

British universities fight back

After a long decade of attrition, British universities may have found a way (and the courage) to stand up to the government that has been their principal tormentor.

THIS year's retreat (to the city of Leeds) by the vice-chancellors (the leaders) of British universities seems to have been rewarding, even daring. The vice-chancellors have resolved to aim at financial autonomy of a kind. Faced with the government's belated but welcome conversion to the belief that a greater proportion of young Britons should follow degree-level courses, and its unwillingness to increase its subvention proportionately, they were even last week flirting with the notion of charging economic tuition fees to all students (see page 375), offsetting the burden on poor students by means of scholarships.

The future of this specific proposal is dubious, but it has brought the issue into the open, and is an excellent means of teasing a government still committed to the view that market forces are a universal emollient. If, the universities may argue, the public water supply and the electricity supply industries are being sold to private owners (see page 372), might not the government also welcome a free market in higher education? The difficulty, as all concerned know well enough, is that the government's political support is already over-sensitive to the cost of higher education. The much more modest proposal, in 1985, that well-to-do parents should pay nominal tuition fees on behalf of student offspring engendered a middle-class outcry and a humiliating retreat. It is unlikely that the same government will accept the vice-chancellors' more radical suggestion when the next general election may be only two years away.

Market forces

Yet something must be done. Part of the explanation for last week's hankering for independence by the universities stems from the realization that there will be no extra funds to cover the cost of the extra 6,000 students recruited to the universities this academic year. The government holds that its decision earlier in the year to double (to £800) the still-nominal tuition fees local authorities are required to pay on behalf of all degree-level students should be sufficient recompense. But the universities are also alarmed that only a kind of engineered chaos can follow from the ambition of the new conduit for public funds, called the Universities (no apostrophe) Funding Council (UFC), to require universities to bid against each other for public money with effect from the beginning of the 1990 academic year. That will

be a market of a kind, but a manipulated market, with the UFC standing proxy for young people seeking an education. Why not — the logic seems to be — create a free market instead?

In the long run, something of that kind is unavoidable. The Secretary of State for Education and Science, Mr John McGregor, was right to remind the universities at Leeds that government subvention of the universities has increased by 9 per cent over the past decade (but student numbers have increased more quickly) and that the collective income of the universities from sources other than the government's higher education budget has increased from 10 to 25 per cent of the total (but much of this is research and contract income, which must be paid for by results). There is certainly no prospect that these trends could sustain an increase of the participation of young Britons in degree-level higher education in all kinds of institutions from 15 per cent (now) to McGregor's still-modest target of 23 per cent by the end of the century.

Scholarships

So how, in the real world, is a transition to autonomy to be arranged? One glaring defect of present arrangements is their uniformity. University salaries (accounting for more than 75 per cent of universities' costs) are nationally negotiated, while tuition fees and maintenance grants for students are fixed by the government. So universities are not genuinely in competition with each other in the true marketplace, that for teachers and that in which students make informed judgements of how they might best (and most economically) get the kind of education they need. The UFC bidding scheme is meant to be a step in that direction, but is hamhanded. The universities' scheme for coupling full economic fees (perhaps five times greater than those now charged) with a scholarship scheme would probably be unworkable, given the need to negotiate with more than 200,000 new students every year. But universities will have no way of measuring their efficiency while students are entirely insulated from the cost of teaching, which is why it would make sense if universities were allowed to charge their students modest fees, perhaps comparable with those now paid indirectly by the government. Among other things, that would test their ability and willingness to run imaginative scholarship schemes, which they could best do individually. □



Cultural chauvinism

Europe and the United States are haggling over airline ownership and television programmes.

EVERYBODY knows that the most cherished symbols of national sovereignty are a national airline and a television broadcasting network. In the 1960s, when independent governments were being created every few months, subsidized airlines and television stations were springing into being at a similar rate. Well-established governments, then, were scornful of this almost ritual hankering after the empty trappings of statehood. Yet it is now apparent that even the most mature governments are not immune from the same self-deception. How else can be explained the smouldering row between the United States and the governments of Western Europe over the ownership of US airlines and, separately, the content of European television broadcasts?

The circumstances are mystifying. Early in the decade, the Reagan administration decreed that US airlines should be deregulated in the sense that those with aircraft meeting specified standards of safety should be allowed to fly whichever routes they chose, charging whatever fares seemed to them to be economic. There followed a brief period during which new US airlines sprang into being and travellers were able to cross North America more cheaply by air than by any other means. But the prospect that competition would provide the smaller airlines with endless opportunities for growth proved hollow; those still surviving have mostly found niches as commuter and feeder airlines, while even the major US airlines have found their mutual competition financially debilitating.

That is partly why Northwest Airlines calculated that it could retain its independence only by seeking foreign investment (eagerly offered by the Netherlands airline KLM) and why the staff of United Airlines has teamed up with British Airways to buy the business from its present owners. But now, to everybody's consternation, the US Department of Transportation has decreed that KLM's proposed stake in Northwest Airlines must be reduced, casting a shadow on the British Airways deal as well. US law requires that no more than a quarter of the stock of US airlines should be owned by foreigners, but the US government has now taken the administrative view that foreigners must not acquire undue influence over US airlines by lending them money either.

The other side of this coin is the restriction of the content of television broadcasts proposed by the European Commission, with the enthusiastic support of several member governments of the European Community. Briefly, the commission proposes that only 15 per cent of the material broadcast by European television stations should come from outside Europe. Mrs Carla Hills, the US Trade Representative, has been in Europe lobbying against this intended regulation on behalf of the US television producers, whose business may be damaged if

Western European television watchers cannot enjoy what has become their usual diet of US programmes.

Both arrangements are unreasonable restraints on trade, and should be openly recognized as such. The time has long since gone when commercial aircraft were essential components of a country's strategic power; to restrict foreign ownership is to restrict competition, thus ensuring that domestic travellers in the United States pay more than would be necessary. The proposed restriction of the foreign content of European television broadcasts will have the same effect; even when people do not have to pay for the privilege of watching television, extra costs will somehow have to be passed on. But is it not worth paying extra to retain Europe's cultural purity? That is the argument that beguiles governments, but can it make sense in a rapidly shrinking world? □

Sale delayed

The postponed timetable for selling Britain's electricity industry will increase uncertainty and reduce the price.

THE plan to sell Britain's electricity industry to private investors was never likely to be smoothly carried through, but the British government's decision last week that its timetable must be postponed by up to six months is a substantial setback, however much ministers may make light of it. There is a political risk; the timetable may now be so stretched out as to confuse the next general election. But there are important issues of principle still to be settled. Among other things, but to nobody's surprise, it is proving difficult to hammer out contractual arrangements between the supply companies, which are due to become retailers of electricity and which will be the first to be sold, and the generating companies, which will own existing power stations, which will be free to build more (as will be the supply companies) and will also enjoy the right to sell direct to industrial consumers of electricity.

That these questions are not simple goes without saying. Although the British government's intention seems to be that the balance between the supply and generating companies should eventually be determined by economics and market forces, the transition from the here-and-now must necessarily be somewhat prescribed. Otherwise, the uncertainties would make the chances of selling either of the two halves of the industry, let alone the National Grid (which is to be owned jointly by the supply companies), at a price reflecting the value of the assets being transferred, small indeed. Even as things stand, and with the government's undertaking to continue owning Britain's first generation of nuclear power stations (and thus to shoulder the substantial decommissioning costs that will eventually arise), the uncertainties are so great that the whole industry is likely to be sold for far less than its present economic value. Some of the eventual owners will profit handsomely, but guessing which they are is, on present showing, mere guesswork. □

Technology growth in French research budget

- Industry gets money and tax benefits
- Aviation and space rewarded

Paris

NUCLEAR technology in France is 'out' and space technology is 'in', if the 1990 civil research and development budget is any guide. In research minister Hubert Curien's second budget of his second term, only the atomic energy centre (CEA), having "reached maturity", has seen its research funding drop. Against an overall civil budget increase of 7.1 per cent to FF45,353 million (\$7,100 million), CEA loses 4 per cent, and for the first time state spending on space, through both the national space research centre (CNES) and contributions to the European Space Agency (ESA), has overtaken spending on nuclear energy research.

France once had a goal of spending 3 per cent of its gross domestic product on research and development by 1992. Last year the figure was 2.34 per cent, and Curien hopes it will rise to 2.38 per cent next year. But if the date has been dropped, the goal remains, and in this budget, industry is being coaxed into spending more on research, with high-technology projects most likely to win government approval.

The government's expansiveness has been offset by slow overall growth in the economy, a call for strict housekeeping and reductions in the growth of spending on defence and nuclear energy research. France is also reducing its high rate of value-added tax, to bring it in line with other European countries before 1992. The loss of tax revenue means that the gamble that the economy will grow through the support of industry and technology must pay off.

Industry will receive a total of FF5,000 million in benefits, partly in the form of extended tax concessions for research and development (FF2,600 million last year). There will be a 30 per cent increase (to FF1,566 million) in the national technology research fund (FRT) which supports emerging areas such as new materials and biotechnology, and through the industry ministry FF843 million, more than three times last year's sum, will go towards technology innovation, such as the European JESSI superconductor programme and the EUREKA high-definition television venture.

Also on the commercial front, this summer's decision to back a fourth-generation of the SPOT remote-sensing satellite has led to FF352 million being set aside next year out of the expected total cost of FF2,500 million. France is a also major

contributor to the ESA space effort: 45 per cent of the cost of the Ariane-5 launcher will come from the French and this, in 1990, will cost them FF1,812 million. The Hermès shuttle and France's contribution to the Columbus space station account for a further FF747 million destined for ESA next year.

A 16.6 per cent increase for research in aviation technology, to FF2,882.6 million, goes mostly to the successful Airbus civil aeroplane, new engines and a helicopter. But the go-ahead has been given for preliminary research into hypersonic passenger flight to develop a competitor to Britain's revolutionary, if neglected, HOTOL space-plane and the West German Sänger equivalent.

Basic research was given less emphasis in Curien's budget presentation but has not been left out. The 'grands organismes', such as the National Centre for Scientific Research (CNRS) and the National Institute for Health and Medical Research (INSERM) between them account for 53 per cent of the civil research budget and in 1990 there will be increases of between 4.6 and 6.9 per cent for CNRS. The laboratory bench receives priority in the raises: salaries and money for equipment go up, and 432 new research posts and 318 technical and support jobs are being created.

The CNRS, the largest state research employer, gets the biggest share, with FF10,331 million and 305 new jobs. The INSERM grant for 1990 will be FF1,825 million, a 5.6 per cent increase, with 88 new jobs. The Pasteur Institutes in Paris, Lille and in French overseas departments will also have 10 per cent more to spend (a total of FF375 million from the state), with the overseas accent on parasitology and research to combat malaria. The environment, particularly research into climatic stability and the ozone layer, will benefit from an extra 13 per cent for the environment ministry and a 30 per cent increase for meteorology.

As in 1989, AIDS research remains "a priority", with a total FF180 million support for the national AIDS research agency (ANRS), in addition to that available within the major state research organizations. About 58 per cent of ANRS grants will go on basic research, 34 per cent on clinical and epidemiological research — including a major survey of sexual habits in France — and the remainder on public health, social and human sciences research. **Peter Coles**

NEWS IN BRIEF

Great le(a)p forward

London

PHYSICISTS at the European particle physics laboratory CERN are jubilant at the success of the first two weeks of experimental runs on the Large Electron-Positron Collider (LEP). With more than 4,000 Z^0 particles detected, they are already preparing their first analyses of this fundamental particle. Before LEP started, only 700 Z^0 's had been detected in all experiments since 1983. Thus, expectations raised at the first trial runs in August, when LEP's first Z^0 was detected after only 15 minutes even though the apparatus was not fully tuned, are being fulfilled. On a good day, 400 Z^0 's can be detected, and further technical improvements should raise this count rate by another factor of 10. By comparison, it is pointed out, the Stanford Linear Collider, previously the leading accelerator in these studies, averages 10 Z^0 's a day.

The LEP data are already sufficient to supersede the SLC measurement of the width of the Z^0 resonance — the range of collision energies over which Z^0 's are produced. Eventually, it is hoped, this parameter will reveal several fundamental features of physics, including the number of neutrino types and the masses of the top quark and Higgs boson. **R.P.**

Exiles' new foundation

Paris

CHINESE student leader Wuer Kaixi and an academic, Yan Jiaqi, were 10 days ago at the centre of the largest meeting of dissidents in exile since the Tiananmen Square massacre on 4 June. Some 167 delegates from different countries — including Hong Kong and Taiwan — and over 200 other participants were present at the Paris Sorbonne to launch a Foundation for Democracy in China. The foundation is headed by Yan Jiaqi and Wuer Kaixi and a 15-member council. Its aim is to form an international focus for opposition to the incumbent Chinese government. **P.C.**

More home comforts

Bangalore

As part of a government effort to persuade Indian scientists working abroad to return home, a scheme called the Transfer of Know-how Through Expatriate Nationals (TOKTEN) has been launched by the Council of Scientific and Industrial Research. Money will be available to pay for qualified expatriates to visit Indian academic institutions and industries, in the hope of generating immediate gains for the host institution as well as promoting links that might draw an expatriate back.

The scheme also aims to find temporary placements for returning scientists and technologists. The creation of new posts is being considered, as are methods for assisting non-resident Indians with the importation of laboratory equipment. **R.R.**

Plutonium generates opposition

Washington

A COALITION of environmentalist and anti-nuclear groups is asking for a legal injunction to prevent the launch on 12 October of the space shuttle Atlantis, whose payload is the Jupiter-bound spacecraft Galileo. A lawsuit filed in the District of Columbia last Thursday claims that the National Aeronautics and Space Administration (NASA) has deliberately underestimated the chances of a launch accident which could release some of the 100 kg of plutonium that Galileo carries to generate power. But those behind the lawsuit would not say whether a reassessment of the risks could even in principle lead to an outcome they would find acceptable, and said that if legal efforts to stop the launch failed they would mount a programme of civil disobedience.

During Galileo's 22-month mission, it will study at close range both Jupiter and several of its satellites, and will drop a probe into the jovian atmosphere. To provide power for its many scientific instruments, it carries two radioisotope thermal generators (RTGs), each containing about 50 kg of plutonium in the form of $^{238}\text{PuO}_2$. The RTGs are not nuclear reactors; radioactive decay of the plutonium dioxide, made up into a number of 5-cm-long fuel pellets sealed into iridium and carbon-boron epoxy-fibre capsules, releases heat which is converted into electric power by thermocouples. Plutonium is used because it emits mostly alpha-radiation, easily captured by the shielding and converted to heat, and little of the more penetrating and more dangerous gamma-radiation.

The lawsuit raises the possibility that if Atlantis, like Challenger almost four years ago, were to explode on launch, the RTGs could be blown apart, scattering plutonium powder across Florida. NASA, on the other hand, says that the RTGs have been physically tested, and shown to remain intact, under conditions more severe than a shuttle explosion.

The groups asking for the injunction are the Florida Coalition for Peace and Justice, an association of local groups active in human rights and anti-nuclear campaigning, the Washington-based Foundation for Economic Trends, led by Jeremy Rifkin, which has lobbied and litigated over many environmental issues, and the Christic Institute. The last of these is an idiosyncratic Washington organization with a history of involvement in anti-nuclear activities, such as the Karen Silkwood case and the radiation leak from the Three Mile Island reactor, but which is also known for its theory that US foreign policy, since the late 1950s, has been steered by a cabal of CIA (Central Intelligence Agency) and ex-CIA operatives of

whom Oliver North is but the latest.

At a press conference to announce the filing of the lawsuit, Bruce Gagnon of the Florida Coalition, Rifkin and Lanny Sinkin of the Christic Institute quoted a variety of different numbers that they said NASA had used at different times for the probabilities of certain kinds of accident. Before the Challenger disaster, they said, the chance of a shuttle explosion had been put at one in 10 million; now it was as high as one in 35. NASA put the chance of plutonium release at one in 2,500, but other official numbers given by Rifkin led to a figure of one in 430. Moreover, it was argued, NASA has suppressed risk assessments that disagreed markedly with its own, and has failed to include all relevant analyses in its Environmental Impact Statement for the Galileo mission.

The lawsuit thus argues that NASA has evaded its responsibilities under the 1982 National Environment Protection Act, and seeks to have the risk assessment analysis re-done, with all facts and figures made public. But, at the press conference, the ultimate purpose of this legal tactic was not clear. Neither Rifkin, Sinkin nor Gagnon would say what probability of plutonium release, determined by a new risk assessment, they would find acceptable. Sinkin likened an Environmental Impact Statement to a risk-benefit analysis; in this case, he suggested, the risk was clearly finite and the benefit, at least to the people of Florida, was nil.

Moving on to larger themes, Sinkin said

PUBLIC HEALTH

Are video terminals dangerous?

Boston

ALMOST four years after it was first proposed, the Mount Sinai School of Medicine in New York has begun a comprehensive, nationwide study intended to provide answers to long-standing questions about the reproductive risk to women office-workers of constant exposure to video display terminals (VDTs). The four-year, \$2-million study began officially last month with funds from the National Institute of Child Health and Human Development.

Michele Marcus, the principal investigator, says the study "will examine the lifestyles, work activities, and health of 8,000 women office workers nationwide". Participants in four cities (New York, Boston, Cleveland and an as yet undetermined city in California) will be assessed by questionnaire, and 800 of the 8,000 will then be monitored closely for one-year during which time the women's menstrual cycles and hormone levels will be monitored, and their urine will be tested for human chorionic gonadotropin as a sensitive

indicator of pregnancy. Provided that some of these 800 women become pregnant during the course of a year, the Mount Sinai study will be the first to monitor women office workers before pregnancy and throughout the first few months. Such a method seeks to control biases that have made questionable the results of some earlier studies. One such study, completed last year at the Kaiser Permanente Medical Center in California, found that women who used VDTs for over 20 hours per week had twice the miscarriage rate of those who did other work.

Louis Slesin, editor of the trade newsletter *Microwave News*, praises the Mount Sinai project and is happy to see the issue receiving federal support. Although the reality of the dangers of VDTs remains highly controversial, Slesin says that the Kaiser Permanente study "made the potential health risks credible", but adds that the Mount Sinai study, when it is completed in 1993, should provide a much fuller picture.

Seth Shulman

David Lindley

MEAT RESEARCH

Third World aid for Bristol?

Tokyo

WHEN a British research institute faced with the prospect of closure by the government recently sent out an international appeal for help, assistance in the form of money intended as aid for developing countries was probably the last thing expected. But that, by way of Japan, is what may come.

In July, A. A. Taylor of the Bristol Laboratory of the Agriculture and Food Research Council (AFRC), formerly the Meat Research Institute, sent a number of letters to colleagues overseas explaining that the British government intended to shut down the laboratory at the end of 1990, and asking for letters of support as a means of persuading the government to keep the laboratory open. One of these letters went to Keiji Umeda, director of the Food Engineering Division of the National Food Research Institute in Tsukuba near Tokyo.

In reply, Umeda suggested that the Bristol Laboratory should apply for Japanese Overseas Development Assistance (ODA) on the grounds that the laboratory could make a good training centre for Africans and Eastern Europeans. Umeda pointed out that the Japanese government announced thousands of millions of dollars in increased aid for

logy and quality assessment, along with basic research on muscle biology. It is considered a center of excellence by meat researchers around the world. But the UK government has decided that the meat industry should pay for the laboratory's research, and is cutting off support.

In his letter to Taylor, Umeda says that "the expertise in meat research at Bristol will never lose its importance for the rest of the world, even if the United Kingdom may not need government research any longer". According to Kiyooki Kato, a colleague of Umeda's at the National Food Research Institute, Umeda spent "very happy days" at the Low Temperature Research Institute in Cambridge in the 1950s, and his suggestion that the Bristol Laboratory apply for ODA money is just "friendly advice".

BRITISH UNIVERSITIES

Who pays for students?

London

GOVERNMENT plans to increase the number of British students in higher education were given a welcome by the Committee of Vice-Chancellors and Principals (CVCP) following their annual meeting at the University of Leeds. But the university leaders warned that expansion without increased Treasury support, as the government proposes, could mean that students would somehow have to pay their own tuition fees.

The need for a new system of student financing emerged as John MacGregor, the Secretary of State for Education and Science, reiterated government plans to increase undergraduate enrolments. He told CVCP there was no doubt that the proportion of students qualified to enter higher education would continue to rise, and expressed delight that universities were responding to the increase in demand "spurred on" by the "incentive" of receiving higher fees from October 1990.

An increase of 6,000 extra students has been reported for entry next year. Moreover, the government has revised upwards its estimate of the Age Participation Index (API, the proportion of individuals aged 18 and 19 going into higher education) which in 1989 stood at 15 per cent: the forecast API for the year 2000 has risen from 18.5 per cent to 23 per cent.

The CVCP welcomed the plans to increase university numbers but in a statement said "it becomes clear the government will not pay for the expansion it desires at a level which will protect high quality". Sir Edward Parkes, chairman of CVCP, foresaw three options by which the increase could be financed: universities could charge fees, a graduation tax

In Britain, the *Guardian* newspaper has suggested that the aid deal could be used by Japan as a "springboard into new markets in Eastern Europe and Africa, in competition with Britain". But Kato denies any such motive, saying Eastern Europe is too poor to import meat from the West or Japan. These nations desperately need to develop domestic production because meat supply is "vital to social stability". And as for Africa, Kato says, Umeda's intention is that the Bristol laboratory should link up with several international agricultural research centres in Africa that are also funded by ODA.

Taylor has contacted the Crown Agents. In a letter to Umeda on 6 September he says that although there does not seem to be much hope, the Crown Agents have requested more details, and adds "perhaps, now that Japanese companies own our motor industry, you can buy the institute".

David Swinbanks



Africa and Eastern Europe at the Western Economic Summit in Paris in July, and that in the 1990s these countries more than any others will be in need of technical guidance and assistance in meat research. But Japan lacks expertise on Africa and Eastern Europe. He suggested that Taylor should contact the Japanese Crown Agents in London, which handle Japanese grant aid, to offer the laboratory's assistance.

The Bristol Laboratory is one of three AFRC laboratories that grew out of the Low Temperature Research Station in Cambridge. It carries out research on meat from slaughter through refrigeration, processing, packaging, on microbio-

related to income could be introduced or the high standards of higher education could be allowed to decline. He suggested that universities could charge full-cost fees, which most undergraduates would be able to meet largely by scholarships derived from government funds.

Such a scheme would enable universities to select potential students from lower economic classes. Universities could play Robin Hood, he argued, by "squeezing the middle classes and employers to pay for the less affluent". Students from lower economic groups would be awarded scholarships based on merit, whereas wealthier students would pay full tuition fees. Universities would publicize the full costs of each course. These would be above those charged to overseas students, which range from £4,000 a year for an arts course to £10,000 for medicine. Parkes emphasized, however, that this system was not a CVCP policy initiative but an avenue to explore.

The CVCP would appeal to present and prospective students, parents and employers for help in persuading the government to provide generous scholarship support. These actions constitute a genuine way forward, Parkes declared, and should not be thought of as alarmist tactics to force the Treasury to give more money to higher education. The vice-chancellors, he concluded, "are not playing games".

The Association of University Teachers (AUT), condemned the CVCP's proposals because they would restrict higher education to the 'privileged few'. Diana Warwick, AUT general secretary, said the CVCP had invited the government to privatize universities and "turn them into yuppie finishing schools".

Ben Webb

Justice not seen to be done?

Washington

A VISION researcher charged by the National Institutes of Health (NIH) with scientific plagiarism, and facing debarment from all federal funds is, not surprisingly, attempting to fight the decision. But he claims that, having been given no real chance to defend himself early on in the proceedings, he may now be the victim of a cursory justice against which he has no legal right of redress.

Two months ago, an NIH investigatory panel declared that C. David Bridges, a vision researcher now at Purdue University, had plagiarized material from a paper sent to him for review, and then published it as his own, in the journal *Science* (see *Nature* 340, 173; 1989). NIH terminated Bridges' current grants and excluded him from service on any peer-review or advisory bodies, actions which the then director of NIH, James B. Wyngaarden, could take by endorsing the panel's recommendations with his signature. But the panel also recommended that Bridges be debarred from receiving federal funds from any source. This NIH have not the power to do, and the recommendation was passed to the Deputy Assistant Secretary for Procurement Assistance and Logistics in the Department of Health and Human Services, who was also to determine the period of debarment. On 18 August, Bridges received a letter from James F. Trickett, DHHS Deputy Assistant Secretary for Management and Acquisition, notifying him that debarment for a period of five years was being considered, and giving him 30 days to respond.

Bridges claims that Trickett's letter give him "no clear idea" of what the next step in the process might be. But Les Ribnik, clerk to George M. Bishop, Bridges' lawyer, said that the letter included copies of the appropriate sections of the Code of Federal Regulations, according to which the debarment procedure is conducted. On 13 September, Bishop wrote to Trickett with a lengthy and detailed response to the charges made by the NIH panel, and requesting a "trial type hearing".

Whatever the next step may be, it will probably take most of the participants into uncharted waters. A "trial type hearing", in which the two sides would be able to cross-examine each other, is not the kind of appeal hearing that would normally be given in such a case. Because there is no constitutional or other elementary legal right to receive federal research funds, suspension of an NIH grant, for example, comes under the jurisdiction of administrative law, not civil law, and a debarment hearing would be conducted somewhat in the manner of a civil hearing in some European countries, in which a

presiding judge asks all the questions and the lawyers for each side are relatively quiet.

But Bridges' complaint is that he has not so far had any opportunity to question the evidence supplied against him, and will not get an opportunity if administrative law takes its course. Indeed, he believes it is entirely within Trickett's discretion whether to grant a hearing on the debarment or not.

Bridges sees himself so far as the victim of dealings that, although entirely legal, are of questionable virtue. When the NIH panel met for two days in Chicago to interview many of the protagonists in the case, Bridges and his lawyer were prevented, despite repeated requests, from having a transcript of the proceedings made, even at their own expense. A preliminary report by the panel was given to Bridges for his response, but nevertheless, he says, he still does not know on what or whose testimony some of the charges were based.

Much of Bridges' rebuttal of the charges against him is an amplified version of his criticisms of the preliminary NIH report which, he says, were not satisfactorily answered in the panel's final version. The NIH review concluded that Bridges could not, as he maintains, have done the crucial experiments before he received a paper from Robert Rando *et al.* for review, because he did not until later acquire sufficient supplies of a tritium-containing reagent, with a restricted shelf-life, essential for the work. Bridges' response sets up a history by which, if earlier supplies of the tritiated reagent were carefully husbanded, and smaller quantities used than Bridges' former technician told the panel, some experiments could have been done before fresh supplies were ordered.

The NIH panel also argues that Bridges suddenly started using the experimental conditions and reagent mixes specified by Rando *et al.*, without any precedent in his earlier laboratory work. Bridges' response is to pick out elements of his previous experiments and draw a thread of reasoning which pulls them together in the later, disputed work.

Bridges maintains that had he been able to make this case to the review panel earlier, and to see and criticize the evidence laid against him, the whole affair would not have got this far. And although he has now assembled an argument which he says would dispel the charge of plagiarism, he fears he will not get the chance to present it. But Trickett, the DHHS official from whom Bridges is waiting to hear, says the right to appeal is one he "insists on", and promises that if Bridges has a case it will be heard.

David Lindley

Squibb builds new laboratory

Paris

THE Squibb Corporation, the US pharmaceutical company, is to spend more than FF280 million (\$43 million) on a basic research initiative in molecular biology at the Université Louis Pasteur in Strasbourg, on the Franco-German border. The deal includes construction of a FF154-million laboratory and support for about 50 scientists for the next seven years. In return, the company will have the right to claim up to half of the projects carried out within the laboratory as theirs, with the option of an exclusive world licence should patentable results emerge.

The 10,000 square metre building should be completed in 1992, on a site in Strasbourg's suburban science park. Under the arrangement with Squibb, if the present 12-year contract is not renewed, the laboratory will remain at the disposal of researchers and after 50 years will revert to the university.

At present, the university's molecular biology laboratory has about 160 staff, mostly employed on permanent contracts by either the state medical research organization, INSERM, or the national science research centre (CNRS). This, says Pierre Chambon, who will direct the new laboratory, means that it is difficult to have a turnover of young researchers. Now, with running costs of around FF18 million per year provided by Squibb for the first seven years, some twenty-five young scientists will be recruited on five-year contracts. To allow this arrangement to work, Squibb has agreed to renew its funding in five-year cycles.

Under the agreement, researchers will remain free to choose their research themes, without interference from Squibb. The work is likely to prove productive for the drugs company, however, as Chambon explains: "the kind of basic research we do, for example on regulatory proteins, will be useful for the design of new drugs over the next 10 to 15 years". The arrangement also does not preclude other companies from taking out licences on results in which Squibb has no interest. Chambon is confident that publication of important results will not be affected by the deal, with a maximum delay of around 45 days in the case of a patent application. Where a licence is taken out, Squibb will pay royalties to CNRS and INSERM.

In recent years, there have been several other examples of pharmaceutical giants, such as Hoechst and Bristol-Myers (now merged with Squibb), investing in basic research. And last July, construction began on a \$32-million neuroscience research laboratory at the University of Oxford under a similar 12-year arrangement with Squibb.

Peter Coles

HYDROELECTRICITY

Danube neighbours' travails

London

THE long-standing dispute between Hungary and Czechoslovakia over the Gabčíkovo-Nagymaros hydroelectric scheme has taken a new turn. Angry at the continuation of the Hungarian moratorium on construction, the Czechoslovak engineers are drawing up contingency plans for "technological countermeasures". These, the Hungarians fear, would pose a far greater ecological threat to northern Hungary than even the original plans.

Last May, in response to growing pressure from 'green' groups in Hungary, the Hungarian government declared a moratorium on its side of the joint project, pending yet another expert survey. Czechoslovakia immediately countered with claims for compensation, saying that without the planned dam downstream at Nagymaros in Hungary to catch its tailwaters, the Gabčíkovo station in Slovakia could never operate at full power. The Hungarians denied these claims; whereupon the Czechoslovak commissioners switched their attention to Hungary's upstream part of the project — the damming of the river-bed at Dunakiliti (where the Danube flows between Hungary and Slovakia), to divert the river into the new Hrusov reservoir which will supply the Gabčíkovo station.

If this reservoir is to be filled by the deadline of December 1990, then the Dunakiliti diversion would have to be completed this October when the hydrological conditions are at their most favourable, with an average flow of about 1,500 cubic metres per second. Once the autumn floodwaters start to rise, the final stage of the diversion operation will have to be postponed, perhaps for a full year, and if the flow exceeds 2,000 cubic metres per second, it may become necessary to destroy the partly finished constructions in order to avoid flooding. Such a postponement, according to Jozef Obložinský, one of the Czechoslovak government commissioners for the project, could have "catastrophic consequences". Already the Gabčíkovo section filling with groundwater, he told the Bratislava daily *Pravda* and if this is not balanced by Danube water entering from above, the foundations of the dam might start to shift.

The projected Czechoslovak "countermeasures" would mean moving the planned diversion to Somorja, some 14 km upstream, where both banks of the Danube lie in Slovak territory. If the Czechoslovak official pronouncements are taken at face value, work will begin on the new diversion "if the two governments fail to agree by October". No agreement has, in fact, been reached, but the Hungarians hope that the announced plans were intended more to strengthen the Czechoslovak

bargaining position than as a real commitment. The proposed new diversion, says László Udvari, the chief Hungarian commissioner for the dam, would have a major impact on the water table in north-west Hungary and would, moreover, have an adverse effect on shipping conditions. Since the Danube is an international river, he says, Czechoslovak "countermeasures" would raise the controversy from a bilateral one to a matter affecting "all those states involved in Danube shipping, and, indeed, all those who have an interest who have an interest in the politics of Europe". The Hungarian government,

Udvari says, still hopes to solve the dispute by "calm consistent negotiations".

So far, however, there seems little likelihood of such an agreement. Czechoslovakia desperately needs the extra energy the Gabčíkovo dam will produce, particularly as its electricity imports from the Soviet Union have been cut for three successive months, and further cuts are expected this winter. The Hungarian government, however, is now committed to go ahead — if at all — only on the basis of a further expert survey and the consensus of public opinion. Even if the Hungarians do, eventually, decide to complete their part of the project, they are not likely to be hurried into meeting the Czechoslovak deadlines. **Vera Rich**

MOTIVE POWER

Powered by hydrogen

Washington

THE hydrogen-powered car pictured here hesitated a little before starting on an unexpectedly cold Washington morning, but then successfully carried Senator Tim Wirth (Democrat, Colorado) and Representative Claudine Schneider (Republican, Rhode Island) on a 50-yard trip in front of the Capitol building. The modified Mercedes car, a research vehicle built by Daimler-Benz in West Germany, was brought over to mark the release of a study commissioned by the Washington-based World Resources Institute that argues for hydrogen, generated by solar electricity, as the environmentally benign transportation fuel of the future.

By burning hydrogen in a fairly conventional engine, the car emits no carbon dioxide or sulphur dioxides, although it inevitably produces some nitrogen oxides from reactions with nitrogen in the air. But if the hydrogen fuel is made by electrolysis of water, using electricity generated by coal or oil, the indirect consequences for the greenhouse effect and for acid rain are as bad as if the car used gasoline. The argument set out in the new report* is that recent efficiency improvements and cost reductions in photovoltaic devices can make electrolysis of hydrogen by solar-powered devices as cost-effective

as other suggested petroleum fuel replacements, such as methanol or ethanol.

The prototype car stores enough hydrogen, in a rechargeable metal hydride absorption device, to take it only about 75 miles.

Basing its calculations on a lighter, more fuel-efficient vehicle, the study estimates that about 280 square feet of solar collecting area would be needed to make the fuel for a car that was driven 10,000 miles in a year. This translates into half a per cent of the entire land area of the United States (or one-sixth of New Mexico) to provide for every car in the country.

Wirth and Schneider, after their short trip, praised the West German government for its support of the development work by Daimler-Benz, and castigated US government and industry for their lack of such research. The sudden drop in crude oil prices of a few years ago was accompanied by a similarly precipitous loss of official interest in energy conservation programmes and alternative energy sources, but alarm over global pollution and warming has begun to revive old concerns.

David Lindley

* *Solar hydrogen: moving beyond fossil fuels* by Joan M. Ogden and Robert H. Williams. World Resources Institute 1989.



The modified Mercedes 230E will reach a speed of 105 mph — but has a driving range of only about 75 miles.

IPA claims harassment

London

THE Soviet Union's new Independent Psychiatrists' Association (IPA) has appealed to psychiatric associations abroad for support following a burglary which it is convinced was deliberate harassment by the Soviet security police. The IPA was founded earlier this year. Its members, from all parts of the Soviet Union, consider that, the Soviet psychiatric establishment is still dominated by those who, during the Brezhnev era, made Soviet psychiatry a byword for political repression and abuse of human rights.

IPA maintains that, in spite of public denials and the visit to Soviet mental hospitals of a "verification" team from the United States, persons of sound mind but inconvenient views are still being committed to Soviet psychiatric hospitals. During the past few months, they have been compiling an archive of known cases as well as offering professional help to patients who, for various reasons, do not wish to seek help through the state system.

Their records, together with correspondence with psychiatrists abroad, were housed in the apartment of one of the association members. A few weeks ago, when the householder was absent for only two hours, the apartment was broken into and the files removed. The break-in, says the IPA, was obviously not the work of an ordinary burglar, since valuables were left untouched.

In an open letter addressed to psychiatric associations throughout the world, Aleksandr Ogorodnikov, a human rights activist and founder-member of the IPA, says that harassment of this kind will not deter IPA, but its members are afraid that seizure of the archives could bring harm to those who have asked IPA for help.

The position of the IPA, Ogorodnikov wrote, would be considerably strengthened if it could become a member of the World Psychiatric Association. Shortly after its foundation, IPA submitted an application for membership, to be considered at the next General Assembly of the WPA, in Athens next week. But the WPA secretariat has refused to process this application, on the grounds that it should have been submitted a full year in advance, when IPA did not exist.

The WPA executive shows signs of proving far less rule-bound in the case of the official Soviet psychiatric society. That organization resigned from the WPA in 1983 to forestall several motions on the agenda for its expulsion. The establishment psychiatrists will also be seeking re-admission to the WPA next week although in its present form that body, too, has existed for less than a year.

Vera Rich

Soviets take French lesson

Vienna

SPEAKING candidly at the 33rd general conference of the International Atomic Energy Agency (IAEA) last week in Vienna, Soviet representatives admitted that the Soviet Union faces serious problems in the continued expansion of its nuclear programme. But, while also announcing a strong commitment to protecting the environment, the Soviet government intends to persist with plans for more nuclear power.

The problem is familiar to Western planners: the Soviet Union would like to increase its power generation capacity by building both nuclear and non-nuclear plants. But nuclear plants run into public opposition, while clean-burning fossil fuel plants are expensive.

With the Chernobyl accident still fresh in the memory, no new Chernobyl-type reactors will be built. Nevertheless, according to Chairman A. N. Protsenko of the State Committee on the Utilization of Atomic Energy, nuclear generating capacity will be doubled by the year 2000, increasing the proportion of Soviet electricity derived from nuclear power from 13 to 15 or 16 per cent.

But "public opinion affects the rate of construction," he said. Protsenko admitted that, while other countries had carried out effective public relations campaigns to make people aware of safety measures at nuclear power plants, "we did not". He attributes what he sees as a "negative attitude" towards nuclear power in the Soviet Union to a "lack of awareness", and revealed that the Soviet Union had been taking lessons from France in how to convince people that nuclear power is safe. A Soviet team visited France one month ago to observe French public relations efforts, and made films for Soviet television.

Slower growth in nuclear capacity has led to plans for more fossil fuel power stations, but the cost penalty of 60 to 70 per cent incurred by equipping a coal-fired power station with scrubbers to remove sulphur and other pollutants has persuaded the Soviet Union to construct its new - and unclean - plants in sparsely populated regions such as the Ural Mountains. This is a "temporary measure," Protsenko added.

Environmentalism has apparently seeped into official Soviet thinking: Protsenko cited a figure of 50,000 million rubles per year (\$79,400 million at the official exchange rate) for the total cost of damage due to environmental pollution in the Soviet Union. One quarter of this is attributed to power generation. Protsenko expects the Supreme Soviet to devote more money to the environment in the 1990 budget. When asked if this would not be difficult in light of the country's prob-

lems in providing items such as basic foodstuffs for its population, he said that it is equally important to provide "foodstuffs and clean air". But Soviet delay in cleaning up the air was criticized by other countries. Rolf Annerberg of the Swedish Ministry of Environment and Energy said that transboundary pollution contributes 75 per cent of the air pollution in Sweden, and called on the Soviets to improve the efficiency of their existing plants instead of building new ones. But Annerberg welcomed the Soviet willingness to discuss the issue. "Before, it was impossible to have anyone say there was a problem", he said.

A number of other Soviet announcements at the conference met with the same mixture of pleasure and scepticism. The Soviet offer to site an international research centre on the consequences of nuclear accidents and nuclear accident management at Chernobyl was repeated. The idea had originally been mentioned to the IAEA board of governors in June.

In a closed meeting, representatives of the IAEA member states that have nuclear power showed some interest in the proposal, although they raised a number of questions that the Soviets chose not to answer. For instance, they asked what research the Soviets had already done at Chernobyl and what purpose an IAEA-sponsored agreement would serve in gaining international access to the site. Although no public response was given, the Soviets were thought to have offered to share results from their investigations and access to waste materials from both the 1986 Chernobyl accident and the 1957 nuclear explosion at Kasli in the Urals (see *Nature* 339, 572; 1989).

The Soviet offer will now be considered by member states before any official response is made. The issue is set to be discussed again at the February board of governors meeting.

IAEA director general Hans Blix of Sweden, whose election to a third four-year term was confirmed by the conference, said that the Soviet Union had also offered to provide a site for an international research project in waste disposal, possibly in the Chernobyl area. But Soviet representatives did not elaborate.

Long-running IAEA issues such as South African and Israeli nuclear capabilities took second place at this year's conference.

South Africa once again promised to consider signing the nuclear non-proliferation treaty (NPT), and will meet with the United States, the Soviet Union and the United Kingdom in December to discuss the matter. The conference postponed for another year a resolution to bar South Africa from the IAEA.

Steven Dickman

Sex statistics unreliable

SIR—Whether or not the British government is right to cancel a proposed social survey of contemporary sexual behaviour, intended to help understand the mechanism of the spread of AIDS, as being “too intrusive” to be supported with public funds, at an estimated cost of £750,000, is perhaps arguable (*Nature* 341, 87; 1989). It would certainly give rise to much salacious comment in the news media, although that would not necessarily have done much harm.

But the real argument against this survey is that its results were bound to have been unreliable. First, although a random sample of people would have been invited to participate, anyone would be free to refuse, so the respondents would be self-selected and probably not representative of the population as a whole. And second, “the questionnaire would have been self-completed in the argot of the social science trade: respondents would have completed the questionnaire themselves, sealing it into an envelope provided, referring to the survey official delivering the questionnaire only in case of ambiguity”.

As this questionnaire, devised by “a large group of British AIDS researchers”, is likely to have been long and complicated and rather technical in parts, it would not have been easily understood by many of those asked to complete it by themselves. Nor can it be assumed that they would always be willing to discuss some of the embarrassing points they did not understand with the survey official. So, quite apart from deliberately false or frivolous answers, a good many of the questions would probably be misunderstood and answered incorrectly, and there would be no possible way of estimating how many of these there were.

The same problems arose with a long and complicated anonymous questionnaire distributed in 1966, designed to produce statistics relevant to the reform of the abortion laws, with demonstrably dubious results (C. B. Goodhart, *Population Studies* 27, 207; 1973).

When trying to decide what should be done, it is a fallacy to suppose that it is better to have bad statistics than no statistics at all.

C. B. GOODHART

Gonville & Caius College,
Cambridge CB2 1TA, UK

Permian of Spiti

SIR—I read with interest the article by A. D. Ahluwalia (*Nature* 341, 13; 1989) under the heading ‘Upper Palaeozoic of Lahul–Spiti’, having just spent 20 days

doing fieldwork near Losar in the Upper Spiti Valley. Ahluwalia’s observations are not correct and are not based on the section exposed in the Savita Nallah. I have worked in the section for the past two years and have made a systematic collection of rocks exposed in the area. Contrary to Ahluwalia’s observation, a complete succession of Ganmachidam section is exposed in the Savita Nallah and has yielded fossils from different strata. The fauna collected from this formation range in age from Middle Carboniferous to Middle Permian. A systematic study of fauna collected from the section is in progress and will be published shortly. The Ganmachidam Formation in this section is overlain by black carbonaceous shales that are rich in fossils and have yielded well-preserved brachiopod fauna including *Lamnimargus himalayensis*, *Spiriferella rajah*. This fauna corresponds to the *Lamnimargus himalayensis* zone widespread throughout the entire stretch of the Himalayas and considered to be of Punjabi age.

ABHINADAN BHARDWAJ

Centre of Advanced Study in Geology,
Panjab University,
Chandigarh, India

Working in Japan

SIR—Although I have yet to visit Japan, I have taught myself rudimentary Nihongo and the most elementary reading and writing in *Kana* (the syllabic ‘shorthand’ Japanese characters) and *Kanji* (the Chinese characters). Thus I feel emboldened to comment on Shahid S. Siddiqui’s question, “Under these attractive conditions, what prevents young US and (*sic*) other European researchers from coming to Japan?” (*Nature* 340, 337; 1989). In answer, Siddiqui mentions most briefly “cultural differences” and “language”.

The Japanese spoken language is admittedly very different from Romance and Teutonic tongues, both in structure and nuance. The written language, however, is virtually incomprehensible without arduous study (as is Chinese). This is the single biggest barrier in the long run.

To visit a foreign country has many benefits, but it can also be a formidable source of anxiety. Very few Americans (or other Westerners) can speak, let alone read or write, Japanese. Thus to live and work in Japan must appear to prospective visitors as virtually impossible.

Few Westerners know that most educated Japanese speak English fairly well (pronunciation excepted). Thus they see themselves faced with an unintelligible tongue in a land where both spoken English, and all the signage are equally unintelligible. It is a far cry from visiting Britain or the Commonwealth, or even the Netherlands or Scandinavia. Most educated Westerners know enough of

other alphabets to cope in Eastern Europe. The Arab countries and Israel, with their seemingly undecipherable alphabets, have considerable English signage, but at least the characters are an alphabet. Consider then the situation for foreigners in Japan where the *Kana* (46 each of two types) represent syllables, not letters, and the *Kanji* (all 1,940 of them) represent complete words.

So, I humbly suggest that Siddiqui misses the main point of his question almost totally. The only steps that could easily be taken on a short-term basis would be crash courses in *Nihongo* in our institutions of higher learning and a change to bilingual Japanese and English signs in public places. The Japanese might also consider their signage in terms of broad use of *Romaji* (Japanese terms in Roman letters) to help accommodate visitors.

In my perhaps limited experience, the cultural differences between the West and Japan are insignificant compared to the language problem.

S. BERLINER III

Heat Systems Incorporated,
1938 New Highway,
Farmingdale, New York 11735, USA

SIR—Several errors crept into the article “Basic research comes alive in Japan” (*Nature* 340, 339; 1989). The KEK laboratory is in Tsukuba; the Ministry of Education, Science and Culture is Monbusho; the new Institute of Advanced Study is being launched by Kyoto and Osaka; and Riken is more naturally not completely capitalized, as it is the common abbreviation of Rikagaku Kenkyujo — the Institute of Physical and Chemical Research — and not an acronym. These small mistakes do aptly illustrate one of the points in the preceding article (“Too few foreign scientists in Japan” *Nature* 340, 337; 1989) — language is indeed one of the problems facing foreign researchers in Japan. However, the English ability of most Japanese scientists and the assistance offered by many fellowships ensure that the problem should not be so large as to discourage those interested in working in Japan.

PHILIP EDWARDS

Institute for Cosmic Ray Research,
University of Tokyo,
Tokyo, 188 Japan

Horrorstruck

SIR—You incorrectly identified Boris Karloff (*Nature* 240, 591; 1989) as being on the left in the picture. He is on the right; Bela Lugosi (seen in other movies as Count Dracula) is on the left.

ALLAN GRANOFF

St Jude’s Children’s Research Hospital,
332 North Lauderdale, PO Box 318,
Memphis, Tennessee 38101-0318, USA

Chapter and verse

SIR—Your reports on recent patents for invention of exceptional importance, and on actions before the British and US courts are extremely useful, but not one of them gives the all-important number of the patent under discussion. Consequently each report necessitates an often lengthy search before the actual patent specification can be studied.

All patent specifications can be found by their distinctive number and thereby easily isolated from the 30 million patent specifications held at the Science Reference Library in London. The search for US patents is made more difficult if the report fails to give not only the number but the full names of the inventors, as the *US Patent Gazette*, issued weekly, does not give the names of the assignees.

I give below, for the benefit of your readers, the numbers of the important patents to which your recent reports have referred:

- (1) "The shadow over immunoassay", (*Nature* **340**, 256; 1989) refers to US Patent 4376110, in the names of Gary S. David and Howard E. Greene.
- (2) "The settlement on AIDS" (**326**, 533; 1987) embraces US Patent 4520113 to Robert C. Gallo *et al.* of 1985 (not 1984).
- (3) The AIDS envelope protein patent (**331**, 649; 1988) is US 4725669 to Myron Essex and Tun-Hou Lee. There is a corresponding European patent application 201588.
- (4) The first-ever animal patent issued in the United States (**332**, 668; 1988 & **338**, 366; 1989) is 4736866 to Philip Leder and Timothy Stewart. The corresponding European patent application is 169672. Each refers to *Nature* **294**, 92–94 (1981) as pertinent prior art, namely the article by F. Costantini and E. Lacy "Introduction of a rabbit β globin gene into the mouse germ line". US Patent Classification List Class 800-1 records about 50 earlier US patents, all directed to multicellular living organisms and unmodified parts thereof.
- (5) The first US patent on tPA for Oxford University (**333**, 383; 1988) is 4751084 to Joseph Feder, William R. Tolbert, Thomas W. Rademacher, Raj B. Parekh and Raymond A. Dwek. The corresponding European patent application is 236289. The Genentech British patent application declared invalid (see Report of Patent Cases 1987 (24) p.553–589) is 2119804 and the corresponding US patent 4766075.
- (6) The US patent involving T7 DNA polymerase-based sequencing technology developed at Harvard by inventors Stanley Tabor and Charles C. Richardson (**340**, 418; 1989) is US 4795699.
- (7) The first US patent for superconductors, to Vander Sande and Gregory J. Yurek (**336**, 607; 1988) is US 4826808. (It was not reported in the *US Patent Gazette*

until 2 May 1989, thereby requiring some 23 weeks of search.)

(8) The US patent for Montagnier, directed to HIV-2 (**340**, 253; 1989) is US 4839288.

FRANK W. COUSINS

43 Emanuel House,
18 Rochester Row,
London SW1P 1BS, UK

Missing formula

SIR—V. A. Huszagh and J.P. Infante's Commentary (*Nature* **338**, 109; 1989) refers to one of the basic problems that inhibit the growth of knowledge in biological sciences: the uncritical accumulation of vast amounts of unconnected data arising from experiments that do not depend on any clearly formulated hypothesis.

Although the authors deserve merit for defining the problem, they do not go far enough in seeking an answer. The examples cited clearly indicate that theoretical solutions in biology are possible only if mathematical methods of another, more basic science fit the problem. Thus, in contrast to physics or even chemistry, biology by itself may be called an incomplete, that is exclusively experimental, science.

Until now, no serious attempts have been made to adapt mathematical logic to biology. Biology therefore lacks both exact principles such as axioms as well as a strictly logical superstructure.

Biologists see the complexity of biology as the only major hindering factor that prevents mathematics from becoming its appropriate theoretical frame. But this impediment might be overcome if there were not a second factor: the participating scientists themselves. Apparently, most of them look upon mathematics as a subject of philosophy — and philosophy is not in their sphere of interest.

MICHAEL OTHMER

University of Mainz,
Department of Pediatrics,
Laboratory of Immunology,
Langenbeckstrasse 1,
D-6500 Mainz 1, FRG

What's in a name?

SIR—Your recent suggestion (*Nature* **341**, 89; 1989) that every member state of the European Communities should have two currencies, its own and the ECU, makes good sense, but it needs one important amendment. Governments for centuries have understood the need to associate bold images with their monetary units. And what image does the ECU suggest? An infinity of small grey men in small cubicles inhabiting the largest grey building in the world.

But what if we name our unit the Einstein? Or, since shortening and initialization is the order of the day, how about 1 Sks for Shakespeare, or 1 Rem for Rembrandt, or 1 Ps for Pasteur, or 1 Vr for Verdi, or 1 Ri for Riemann?

One could believe that a decision over which particular name to choose would cause great argument, sufficient to disrupt more than one European summit. But the need also to name the various multiples of the primary unit would accommodate quite a few great names. And for the primary unit itself there is one special name that might well recommend itself to everybody, partly from a sense of irony and partly as a much-needed retribution. So let the primary unit be named after someone whom the world of his own day buried in a pauper's grave. Let it be 1 Mz for Mozart.

A sensibly named currency would of course soon drive out national currencies on a European scale. But would it not also drive out other currencies worldwide? For who would want to carry dull dollars in their wallets when they could be carrying Mozarts? I ask you.

FRED HOYLE

102 Admiral's Walk,
West Cliff,
Bournemouth BH25 1TF, UK

SIR—In your leading article entitled "Funny money" (*Nature* **340**, 580; 1989) you did not address the matter of its physical form. I suggest that consideration be given to replacing paper money by plastic or plasticized paper, of a size and shape like those of credit cards but thinner.

Cards should be washable, of standard weight and colour-coded. (Perhaps US notes are the only ones that still lack this feature.) They could have various metal strips to make them easily acceptable, distinguishable, sortable and countable in machines (for parking meters, telephone boxes, bus/train tickets, change, and so on). Suitably disposed fenestrae or notches could help the blind as well as the machines. Incorporation of specific diatom shells could add an unforgeable forensic feature, as could fluorescent threads.

Denominations could be variously printed for different countries (for example, red: 1 Eurobuck = DM2 = FFfr2 = 1,000 IL) on one side; on the reverse, the international value would be indicated in Esperanto. There would be no religious symbols or slogans in any text. If and when necessary, changes could be embossed.

Paper money is dirty and ephemeral: it is time we evolved better things.

RALPH A. LEWIN

Scripps Institution of Oceanography,
University of California, San Diego,
La Jolla,
California 92093, USA

Model lung breathes true

An interesting calculation of how the human lung functions has appeared in an Italian journal whose fortunes appear (from inspection) to be reviving.

ONCE upon a time, in the 1950s, *Il Nuovo Cimento* shared with the Japanese journal *Progress in Theoretical Physics* the distinction of being among the most eagerly read in its field, but it was then eclipsed by the emergence of rapid-publication journals such as *Physical Review Letters* and, later, of *Physics Letters*.

But *Il Nuovo Cimento* seems to be perking up. Nowadays, it is almost always possible to find something of interest in an issue. Part of the explanation is that, for the past eleven years, there have been four sections dealing separately with nuclear physics, astrophysics, condensed matter and general physics, which means that there are in principle four issues a month. In some sections, the publication lag (one of the reasons for the original decline) is still uncomfortably long — some papers seem to take a year in the process of publication. But the more obscure backwaters of general relativity on which the journal retained a lien throughout its bleak period are no longer the staple diet for *Il Nuovo Cimento*'s faithful readers. Last month's condensed matter issue, for example, contains an intriguing calculation of gas transport in the human lung of a kind that would normally have been found in a physiological journal.

As F. Bonsignori and M. Salvini, both physicists from the University of Bologna, describe it, the problem is really fairly straightforward (*Il Nov. Cim.* **11D**, 1247; 1989). The trick, which has been used before, is to represent the lung and its airways, the trachea and bronchi, by a kind of one-dimensional model. They follow physiologists in supposing that there are 23 generations of successive refinement in the lung, at each of which the total area of the walls across which gas can be exchanged increases. The problem then becomes that of calculating the processes by which exchangeable gases reach the walls of the lung and diffuse through them.

The physiological data used to ground the model are striking in themselves. The cross-section of the airways increases from 2.54 cm² in the trachea to 11,800 cm² at the finest and twenty-third subdivision of the alveoli, whose total volume is a third of the capacity of the lung as a whole (taken to be just over 1.5 litres, everything included). The most obvious consequence of this state of affairs is that gas velocities progressively decrease from the mouth to the alveoli, so that in the main airways into

the lung, turbulent motion and convection are chiefly responsible for gas transport, but molecular diffusion dominates in the alveoli.

Bonsignori and Salvini acknowledge that their model is not original, but they do claim to have improved on earlier calculations (in the 1970s) by improving on some of the approximations on which previous calculations have been subject. For one thing, they reckon to have used a better representation of the cross section of the lung with distance along the airways. More pointedly, they allow for the way in which the effective diffusion constant can vary with the velocity of the flow and with place along the airways. In the alveoli proper, the implicit assumption is that the gases in the lung exchange their contents freely with the blood, which is for practical purposes regarded as an infinite reservoir.

The objective is to write down and then to solve a differential equation for the diffusion and convection of gases along the airways; Bonsignori and Salvini judiciously discuss the boundary conditions that should apply at the walls of the airways, acknowledging that in the alveoli, at least, there are grounds for believing that the concentration of one component of gas inspired will be virtually constant along directions perpendicular to the walls of the alveoli, and that from some point sufficiently deeply buried in the network, the concentration of all components will be virtually unchanging in space and time. But this implies that there are two formally different sets of boundary conditions that may be relevant, whence the suggestion (not further followed up) that some mixture of the two would be most appropriate.

The variations of the effective diffusion coefficient, which determines the effectiveness of mixing under different circumstances, are tackled semi-empirically. In the entering airways, mixing is much more efficient than would be determined by molecular diffusion, but in the alveoli, molecular diffusion is king. Part of the construction of the equations to be solved entails fixing parameters by reference to the known parameters of the lung, using the helpful (and reasonable) conditions that air should not accumulate anywhere in the system. The remarkable feature of the calculation is that it leads to an explicit result, although one in which several numerical approximations have been

made.

The outcome of the calculation is much what physiologists would expect. The authors restrict themselves to the case in which it is supposed that somebody takes in a single substantial breath, holds it for a little while and then empties the lungs. What the solution of the equation shows is, first, that a component of the inspired gas progressively penetrates through the system with the passage of time, and that the process is the more rapid as the incoming flow is faster.

The supposition that, in the outer reaches of the system, the concentration of particular components of the inspired gas quickly become more or less constant readily falls out from the calculations, but concentrations in the inner reaches quickly fall to zero as gas is rapidly absorbed across the alveolar walls. The time constant of the system appears to be a few seconds — rather more than 4 seconds in one significant context — which accords with general experience of trying to hold one's breath. It may or may not be coincidence, but the calculations appear to show that, when breathing air, the model lung reaches in just over a second a condition in which the oxygen concentration is virtually constant throughout the inner part of the respiratory system, taken to be the last seven generations of airway refinement.

Inevitably, the authors promise to produce a second set of more refined calculations, paying particular attention to the vexed question of the boundary conditions most appropriate in different parts of the respiratory system. At this stage, however, nothing is said about the assumptions that the gas in the lungs is incompressible, or that there is no interference between the different components of the inspired gas that people breathe. No doubt the next step will be to collect some carefully controlled experimental data in the hope of matching calculation and experiment.

The interest of all this is not simply that it is what seems to be a reasonably realistic calculation of the properties of a complicated physiological system, but that it appears in a straightforward physics journal. No doubt the bibliographic retrievers will alert physiologists to what they may otherwise overlook, but that it should have appeared there at all is a neat sign of how broadly catholic is the Italian Physical Society.

John Maddox

Genes and segmentation

Julian Lewis

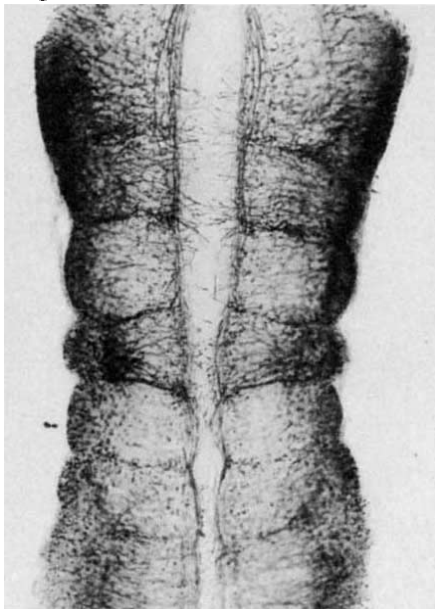
SIXTY-SIX years ago, in the early days of *Drosophila* genetics, the first of a strange class of mutations was described¹. The mutant fly, called *bithorax*, had its third thoracic segment, which should normally carry a pair of balancers, partially converted into a duplicate of the second thoracic segment, which carries wings: the relationship between location in the body and selection of a developmental pathway had been upset. This was the beginning of a chain of discoveries that began slowly but is now taking us at a dizzying pace towards a revolution in our understanding of the mechanisms of development. Findings from *Drosophila* are illuminating the development of other animals too. From several recent papers, including one on page 405 of this issue², it appears that there are profound similarities between vertebrates and insects in the genesis of their body plan, in particular with respect to segmentation and the genes that give each body segment the character appropriate to its location.

Mutations that, like *bithorax*, cause segments of the body to develop as though misinformed about their location define a class of genes called homoeotic selector genes. As more such genes were discovered in *Drosophila*, it became apparent that many of them were grouped into one or other of two tight clusters in the genome, called the antennapedia complex (ANT-C) and the bithorax complex (BX-C). ANT-C contains genes specifying the differences between anterior segments, while BX-C is responsible for the pattern of differences between posterior segments. If, for example, the whole of BX-C is deleted, all segments posterior to the mid-thoracic segment³. (For simplicity, I use the term 'segment' loosely; in the technical jargon, 'parasegments' would be the more exact term here.)

More detailed analysis of BX-C revealed a curious regularity. Mutations at different loci within the complex affect different domains of the body, and the order in which these loci are arranged within BX-C corresponds with the spatial ordering of the affected domains along the body axis³. In other words, as one passes to more posterior levels in the body, the genetic elements of BX-C are called into play sequentially according to their location in the chromosome. The same is true for ANT-C in *Drosophila*⁴. And in the beetle *Tribolium*, where a single giant complex combines the functions of ANT-C and BX-C, the correspondence between the genetic map and the body axis is preserved over its whole length⁵.

Borrowing terminology coined for the beetle, it is convenient to call any cluster of homoeotic selector genes organized in this fashion a HOM complex and its members HOM genes⁶.

HOM complexes have a further property that seems to be crucial for their function. Cells of each body region in *Drosophila* become stamped during development with a permanent record of their location along the head-to-tail axis. For example, when imaginal disks (the immature larval rudiments of adult structures) are transplanted, their cells persist in differentiating according to their original location⁷. However, removal of a



Rhombomeres in a chick embryo hindbrain (from ref. 21) — homology running deep?

HOM gene (by X-irradiation) from a small clone of cells within one imaginal disk late in the development of the animal can convert that group of cells to the character appropriate to another region⁸. This suggests that a cell's memory of its antero-posterior location in the embryo depends on maintenance of the extent of activation of its HOM complexes, which presumably must become frozen at the level to which they have been driven by some transient embryonic signal, graded along the body axis. (There are intriguing parallels here with the phenomena of X-inactivation and position-effect variegation⁹, where again a change in gene expressibility spreads to a variable extent along the chromosome, becomes frozen at a critical time in development, and is thereafter remembered by the cell and its progeny.)

The meaning of the 'extent of activation' of a HOM complex has been clarified by molecular cloning of BX-C. This

complex, though it governs the patterning of at least nine segments, comprises only three protein-coding genes, *Ubx*, *abdA* and *AbdB*. *In situ* hybridization shows that these genes are expressed serially along the body axis in accordance with their chromosomal locations¹⁰. However, the genes are not expressed uniformly in all the cells of a given segment; instead, the gene products are used in a more subtle and variegated way to give each segment a unique character. Each of the structural genes is regulated by a complicated entourage of enhancers, and its expression is differently modulated within each of the segments where it is required for normal development. Many of the homoeotic mutations disturb the enhancers rather than the structural genes, and the correspondence between chromosomal location and spatial domain in the body holds true in a more complete and detailed way for the regulatory loci than for the structural genes themselves. It seems the enhancers become 'open for business' — that is, receptive to *trans*-acting transcription factors — sequentially in successive segments according to their position in the chromosome, thereby exposing the associated structural gene to a particular segment-specific style of regulation¹¹.

It is not yet clear what drives the progressive opening-up of the HOM complexes along the body axis (though information that should lead to an answer is rapidly accumulating¹²⁻¹⁴); nor is it known how the stage of HOM activation attained during embryogenesis is preserved and inherited from parent cell to daughter cell thereafter. But whatever the molecular mechanisms may be, the properties of the HOM complex make it a very useful developmental gadget: the extent of its activation can represent for the cell a record of the magnitude of an influence to which it or its ancestor was exposed in the embryo. The 'magnitude' of the influence might mean its duration, or its intensity, or some product of the two: one could conceive of a HOM complex being used to record the value of almost any sort of graded developmental variable. Indeed, recent work suggests that in vertebrates HOM complexes are used in a variety of ways to specify the pattern of several different parts of the body — not only the axial structures but also the limbs^{15,16}.

HOM and Hox

It would not be possible to speak of HOM complexes in vertebrates had it not been for another discovery in *Drosophila*. All of the fly's HOM genes were found to include a sequence of 180 base pairs, the famous homoeobox, coding for a highly conserved DNA-binding protein domain. Evidently all these genes, both in BX-C and in ANT-C, have arisen by duplication and divergence from a common ancestral

homoeobox-containing gene, whose product presumably acted as a master-regulator of the activities of other genes. The discovery of the homoeobox made it possible to fish for homoeobox-containing genes elsewhere in the *Drosophila* genome and in the genomes of other animals.

In mammals and other vertebrates one finds, as in *Drosophila*, many genes that contain homoeoboxes, and among them a special subset that, by virtue of an astonishing set of parallels with the insect HOM genes, may also be called HOM genes. They are grouped into a small number of tightly linked clusters, or HOM complexes. In the mouse, at least four such complexes, called Hox-1, Hox-2, Hox-3 and Hox-5, have been discovered. The individual genes of a given HOM complex, as in *Drosophila*, are expressed in different regions along the head-to-tail axis of the body, the pattern being clearest in the central nervous system¹⁷.

Moreover, several of the mammalian HOM genes can be identified as individually homologous to specific *Drosophila* HOM genes: thus *Hox-2.1* is a mouse homologue of *Drosophila Scr*, being more closely related to it than to its chromosomal neighbours, *Hox-2.2* and *Hox-2.6*. Similarly, *Hox-2.6* is a specific homologue of *Drosophila Dfd* — and so on. Again, a simple correlation is seen between the location of a gene within its complex on the chromosome and the region where it is expressed in the body of the animal: the further the gene lies from the beginning of the HOM complex, the more posterior the domain to which its expression is restricted^{18,19}.

The similarity of organization and the pattern of homology and divergence in the HOM complexes of insects and mammals suggest that their common ancestor already possessed a HOM complex containing counterparts of at least some of the insect HOM genes, and used this complex as insects (and probably we) do, to specify the differences between successive regions along the head-to-tail axis of the body. The homology of the HOM genes between insects and vertebrates suggests also a homology of body parts, defined according to the HOM genes they express. But the drawing of parallels between insects and vertebrates immediately raises problems, for at first sight their bodies seem to be constructed on radically different principles.

The insect body is built from segments — a series of modules that are repeated with variations along the anterior-posterior axis. Differences of HOM-gene expression specify the differences between the segments, and the boundaries of HOM-gene expression domains are exactly aligned with segmental (or para-segmental) boundaries, as a result of regulation of HOM-gene expression in the

early embryo by the genes that generate the segments¹³.

What then is the vertebrate counterpart of an insect segment? Are vertebrates segmented in the same sense in which an insect is segmented? Initial attempts to answer this question naturally enough focused on the segmentation that is most obvious in our own bodies, corresponding to the series of somites in the embryo or of vertebrae in the adult. But it has been persuasively argued²⁰ that this could not be homologous to the segmentation of the insect: the somitic segments of a vertebrate develop in a way that is completely different from the body segments of an insect and, according to most accounts, they must have evolved after the vertebrate and insect lineages diverged.

Rhombomeres

Several recent papers cast a new light on this problem and suggest that we have all been barking up the wrong tree: in the search for homology, we have to focus on another, and more ancient, system of segments in the vertebrate body. The expression domains of many of the mammalian HOM genes in the central nervous system have their anterior boundaries in the hindbrain. When examined at the appropriate stage of embryonic development, the hindbrain consists of a series of seven or eight bulges known as rhombomeres, delimited by slight constrictions (see figure). This obscure detail of embryonic anatomy took on a new significance when it was shown²¹ (in the chick) that the bulges and constrictions reflect a regular repetitive segmental organization of the neurons and glia in the neural tube, correlated with the series of cranial motor-nerve roots and cranial sensory ganglia, and thereby correlated also with the series of branchial arches — or gill arches — that they innervate. *In situ* hybridization studies showed that rhombomeres are also delimited biochemically: the *Krox-20* gene is expressed in rhombomeres 3 and 5 (ref. 22), whereas the *int-2* proto-oncogene is expressed (transiently) in rhombomeres 5 and 6 (ref. 23).

The rhombomeres thus define a set of landmarks in the hindbrain that must have some major developmental significance. A recent paper in *Nature*²⁴, and another in this week's issue², now reveal that the rhombomere boundaries correlate precisely with the expression boundaries of HOM genes. For example, *Hox-2.9* is expressed in rhombomere 4, in a region delimited so sharply that one is reminded of the compartment boundaries seen in *Drosophila*. The other four genes of the Hox-2 complex that have been examined in this way (*Hox-2.1*, *-2.6*, *-2.7* and *-2.8*) are also expressed in domains, the anterior boundaries of which coincide with the boundaries of rhombomeres. What is more, for each step from one of these

genes to the next along the chromosome, the anterior expression boundary jumps two rhombomeres closer towards the head, just as the anterior boundaries of the expression domains of successive *Drosophila* HOM genes are, for the most part, spaced at two-segment intervals²⁴.

Perhaps these parallels between the rhombomeres of the vertebrate and the segments of the insect are just a trick of convergent evolution — a case where two classes of organisms have independently evolved systems of segments, which would thus be analogous but not homologous, and independently coupled the activation of a HOM complex to the segmentation machinery with the same spacing of two segments per HOM gene. Alternatively, there may truly be a homology which runs deep.

Comparative anatomy and palaeontology suggest that the series of gills is a more ancient feature of the chordates than is the series of somites^{25,26}. It has been conserved faithfully in vertebrate embryos, even though subsequently modified out of all recognition in the adult mammalian body, where it gives rise to much of the face and neck. This conservation of the system of embryonic gill rudiments was emphasized by the founding fathers of evolutionary theory as evidence for the common ancestry of vertebrates. It would be pleasing if the same structures were to serve in our own day as a key to the relationships between vertebrates and invertebrates, and as an exemplar of the universality of the mechanisms of animal development. □

Julian Lewis is in the ICRF Developmental Biology Unit, Department of Zoology, Oxford OX1 3PS, UK.

1. Bridges, C.B. & Morgan, T.H. *Carnegie Inst. Wash. Publ. no. 327* **137**, 152 (1923).
2. Wilkinson, D.G., Bhatt, S., Cook, M., Boncinelli, E. & Krumlauf, R. *Nature* **341**, 405–409 (1989).
3. Lewis, E.B. *Nature* **276**, 565–570 (1978).
4. Mlodzik, M., Fjose, A. & Gehring, W.J. *EMBO J.* **7**, 2569–2578 (1988).
5. Beeman, R.W. *Nature* **327**, 247–249 (1987).
6. Akam, M. *Cell* **57**, 347–349 (1989).
7. Hadorn, E. *Scient. Am.* **219**, 110–120 (1968).
8. Morata, G. & Garcia-Bellido, A. *Roux's Arch.* **179**, 125–143 (1976).
9. Baker, W.K. *Adv. Genet.* **14**, 133–169 (1968).
10. Akam, M. *Development* **101**, 1–22 (1987).
11. Peifer, M., Karch, F. & Bender, W. *Genes Dev.* **1**, 891–898 (1987).
12. Driever, W. & Nüsslein-Volhard, C. *Cell* **54**, 83 (1988).
13. Ingham, P. *Nature* **335**, 25–34 (1988).
14. Ruiz y Altaba, A. & Melton, D.A. *Nature* **341**, 33 (1989).
15. Oliver, G. *et al. Genes Dev.* **3**, 641–650 (1989).
16. Dollé, P., Izpisua-Belmonte, J.-C., Falkenstein, H., Renucci, A. & Duboule, D. *Nature*, (in the press).
17. Holland, P.W.H. & Hogan, B.L.M. *Genes Dev.* **2**, 773–782 (1988).
18. Graham, A., Papalopulu, N. & Krumlauf, R. *Cell* **57**, 367–378 (1989).
19. Duboule, D. & Dollé, P. *EMBO J.* **8**, 1497–1505 (1989).
20. Hogan, B., Holland, P. & Schofield, P. *Trends Genet.* **1**, 67–74 (1985).
21. Lumsden, A. & Keynes, R. *Nature* **337**, 424–428 (1989).
22. Wilkinson, D.G. *et al. Nature* **337**, 461–464 (1989).
23. Wilkinson, D.G. *et al. EMBO J.* **7**, 691–695 (1988).
24. Murphy, P., Davidson, D.R. & Hill, R.E. *Nature* **341**, 156–159 (1989).
25. Romer, A.S. *The Vertebrate Body* 5th edn (Saunders, Philadelphia, 1976).
26. Jefferies, R.P.S. *The Ancestry of the Vertebrates* (British Museum Natural History, London, 1986).

New clue to Ras lipid glue

Douglas R. Lowy and Berthe M. Willumsen

NORMAL Ras proteins encoded by the *ras* genes and oncogenic Ras proteins encoded by mutated *ras* genes (oncogenes) are synthesized as a precursor in the cell cytoplasm and then modified and translocated to the cytoplasmic face of the cell's outer membrane¹. Localization to the membrane is critical for efficient Ras function, as the biological activity of those mutant Ras proteins that remain in the cytosol is severely impaired². Until very recently, most investigators believed that it was the attachment of palmitate near the carboxy-terminus of Ras precursors that mediated their movement to the plasma membrane³. Now, evidence from three laboratories⁴⁻⁶, each using a distinct approach and making a unique contribution, indicates that the lipid invariably added to mature Ras proteins is a farnesyl isoprenoid, and that palmitate is only an optional extra. The farnesyl moiety common to all mature Ras proteins is derived from farnesol, which is a branched-chain, polyunsaturated hydrocarbon alcohol intermediate of sterol biosynthesis. As farnesol is derived from mevalonate, pharmacologic agents which interfere with 3-hydroxy-3-methylglutaryl coenzyme A (HMG CoA) reductase (the enzyme that catalyses mevalonate synthesis), or other steps before farnesol synthesis, may be able to inhibit membrane localization of Ras protein, thus providing a potential approach to interfere with Ras-protein function.

The critical clues leading to these findings came from the recognition of the similarities between the post-translational modifications of fungal mating factors, such as yeast α -factor^{7,8} and those of Ras. Indeed, a yeast mutant defective in post-translational modification of mating factor seems to cause a similar defect in yeast Ras^{5,8}. The last four amino acids of both α -factor and Ras (see figure), when first translated, can be described as Cys-A-A-X, where Cys is cysteine, A is an aliphatic amino acid (most often the hydrophobic amino acids leucine, isoleucine or valine), and X is any amino acid³. The carboxy-terminus of the mating factor has been shown to undergo several post-translational changes, including proteolytic cleavage of the A-A-X, carboxymethylation of the (now carboxy-terminal) cysteine and its farnesylation via thioether linkage to its sulphur atom⁷. Virtually identical modifications of Ras have now been shown to occur^{4-6,9,10}. Furthermore, a Cys-A-A-X box, when placed at the carboxy-terminus of a protein that does not normally have one, is sufficient to direct its binding to an isoprenoid (presumably a farnesyl moiety)⁴.

Palmitate, which was formerly believed to be linked to the cysteine residue in the Cys-A-A-X box, has now been shown to be linked instead to cysteine residues located in the first few amino acids upstream from the box. Moreover, the palmitoylation somehow depends upon post-translational modification of the box^{4,4}. Rather than representing an obligatory step in Ras-protein function, palmitoylation can now be viewed as a modification that may modulate the biological activity of a given Ras protein. Supporting this hypothesis, palmitoylated Ras proteins have been found to be more firmly anchored in the membrane than non-palmitoylated versions and to be more active in stringent bioassays⁴. Most eukaryotes, including mammals, contain multiple *ras* genes that encode proteins with considerable sequence divergence in the 20 amino acids immediately upstream from their Cys-A-A-X box¹. This sequence divergence is associated with variation in the precise location and number of cysteine residues that can be Palmitoylated (see figure). As the biological activity of a Ras protein seems to vary with the amount of palmitate that is added, each Ras protein may have its own particular level of activity, presumably to allow distinct functional roles to be better served than they could be by a

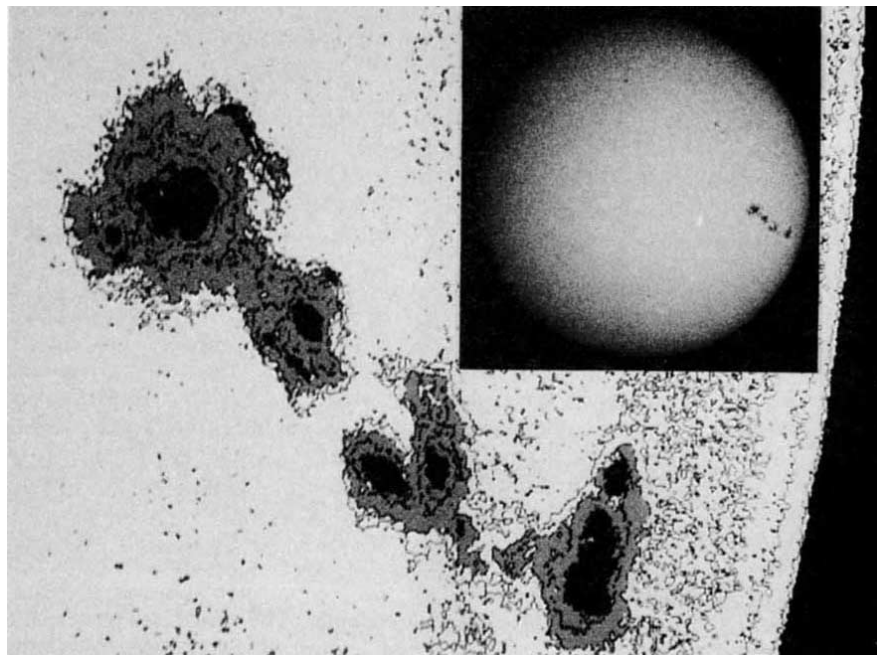
<i>ras</i> ^H	G	P	G	Ⓢ	M	S	Ⓢ	K	Ⓢ	C	V	L	S
<i>ras</i> ^N	T	N	G	Ⓢ	M	G	L	P	C	V	V	M	
<i>ras</i> ^K	P	G	Ⓢ	V	K	I	K	K	C	I	I	M	
<i>ras</i> ^K	K	K	K	K	S	L	T	K	Ⓢ	C	V	I	M

Carboxy-terminal 12 amino acids encoded by the three mammalian *ras* genes: the last four residues constitute the Cys-A-A-X box. The farnesylated cysteine residues in the Cys-A-A-X sequence are boxed; upstream cysteines, which may be palmitoylated, are circled. The two termini shown for *ras*^K are the result of alternative splicing of two exons (A and B).

single Ras protein. In addition, the palmitate bound to Ras has a much higher turnover rate than does the protein itself, which raises the possibility that alterations in the turnover rate of bound palmitate could represent an additional level of regulation¹¹.

Other genes have also been identified that encode proteins with a terminal Cys-A-A-X box, with or without upstream cysteines that may be palmitoylated. Many of these proteins bind guanine nucleotides, as does Ras, and have additional regions of sequence similarity to Ras. But others, such as the nuclear lamins or HMG1 reductase, are similar to Ras only by terminating in the Cys-A-A-X sequence. It seems likely that at least some of these proteins will also be farnesylated. They could represent some of the as yet unidentified proteins that have been shown to incorporate mevalonate^{6,12}, as farnesol is derived from mevalonate. As lamins are associated with

Giant sunspot clusters mark solar maximum



With only 3 months until the solar maximum, violent activity — flares and sunspots — at the Sun's surface is increasing. This massive cluster of sunspots, photographed recently by N.W. Scott, traversed the Sun's face in 10–12 days. (Note: binoculars and telescopes should not be used for direct viewing of the Sun.) □

nuclear membranes and HMG1 reductase with the endoplasmic reticulum, their isoprenoid binding¹² suggests that this modification facilitates general membrane association, with other factors determining the particular membrane to which the protein is bound.

The number of proteins that are modified in the same way as Ras may be even greater than expected because the presence of a Cys-A-A-X box may be too stringent a criterion for their identification. Whereas the cysteine is obligatory as the first amino acid of the box in Ras proteins, the second or third positions can be occupied by some non-aliphatic amino acids (lysine, threonine, tryptophan) without affecting the post-translational modification or biological activity of the proteins^{1,2,9}. Thus, the Cys-A-A-X box may have more the character of a consensus sequence, which may be present in many proteins.

As mutant, highly transforming *ras* genes are found in many tumours, inhibition of Ras function would seem to be an eminently worthwhile goal. Whereas past studies have already implied that it might in theory be inhibited through preventing the post-translational modifications required for membrane binding, the connection between the sterol biosynthetic pathway and Ras processing means that such inhibition can be tested experimentally, using known inhibitors of the pathway. Indeed, two related inhibitors of mevalonate synthesis, compactin and mevillin, have been shown to inhibit isoprenylation of Ras and to prevent its membrane association^{4,5}. There are several reports that such inhibitors, when used at concentrations much higher than those required to inhibit most cholesterol synthesis¹³, can inhibit cell growth *in vitro* and sometimes induce cytotoxicity. Under some growth conditions, rodent cells transformed by the T24 *ras* oncogene may be more susceptible than the parental cells to growth inhibition by mevillin¹⁴.

But several theoretical problems underlie the potential clinical application of this approach, which, to be successful, must inhibit tumour growth to a much greater

extent than it interferes with normal cell physiology. Although the inhibition of sterol synthesis might, in theory, be overcome by providing exogenous cholesterol (and/or its derivatives), it is not easy to imagine how it could be at least relatively selective for transforming Ras proteins when normal Ras and perhaps numerous other proteins are farnesylated, presumably by the same enzymatic process. Therefore, despite the understandable

excitement generated by contemplation of this approach, it still remains to be demonstrated whether inhibitors of mevalonate synthesis are of use in the treatment of tumours. □

Douglas R. Lowy is in the Laboratory of Cellular Oncology, National Cancer Institute, Bethesda, Maryland 20892, USA and Berthe M. Willumsen is at University Institute of Microbiology, DK-12353 Copenhagen, Denmark.

RADIO GALAXIES

Lining up the evidence

John Hutchings

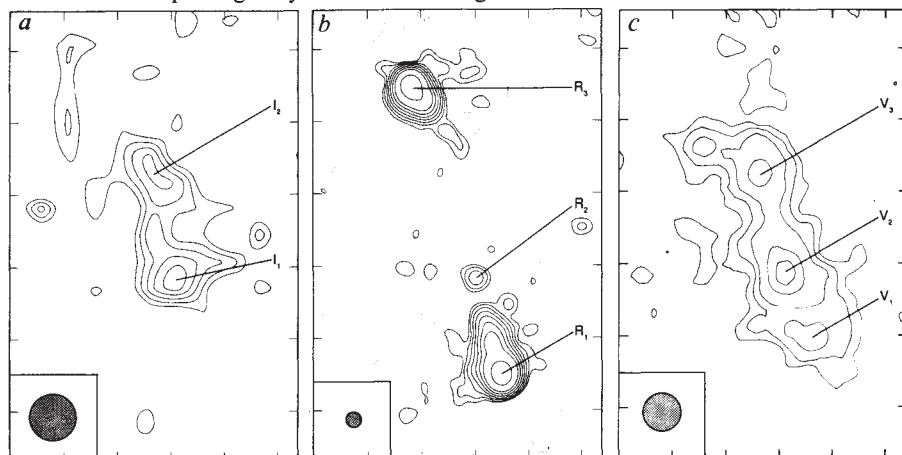
THAT the radio images of radio galaxies and quasars seem to be very different from their optical images has long been a cause for debate. The general view is that the radio structure is determined by the energetic events occurring in the galactic nucleus which are largely unrelated to the distribution of the stars and dust that determines the optical structure. The typical bipolar form of the radio emission probably originates in the angular momentum of the central engine, and is probably aligned with the engine's spin axis. Naively, it used to be supposed that this axis, with little evidence, would be parallel to that of the visible galaxy, so that the disk or ellipse of stars should be perpendicular to the radio lobes. New evidence, however, seems to contradict this expectation, prompting Rees to propose a new model of the mechanisms involved¹.

The evidence comes from large new samples of galaxies that have been mapped with arcsecond resolution or better at both optical and radio wavelengths. Baum and Heckman² find in a sample of nearby (low-redshift) luminous radio sources that the 'orthogonal' rule applies only for radio structures that greatly exceed the host galaxy in size. In most cases the optical galaxy and the radio

structure are disturbed and comparable in size. However, the authors also find that there is a strong correlation between the ionized gas (as distinct from the stars) in the galaxy and the radio structure, both in total power and spatial extent. They conclude that either the ionizing radiation emerging from the nucleus is collimated (or beamed) or the radio-emitting material itself is significantly disturbing the structure of the host galaxy. Independent detailed studies of several active galaxies have indicated that both phenomena occur fairly commonly.

At greater distances (higher redshifts), the situation is different. Beyond a redshift of 0.6, McCarthy *et al.*³ and Stockton and Lilly⁴ find an alignment of radio and optical structure. For redshifts over 1.0, the alignment becomes tighter (within 15°) and is more common, occurring in more than 85 per cent of cases (see figure). In addition, ionized gas can be seen at all places in — or beyond — the radio structure, even at hundreds of kiloparsecs outside the stellar boundaries of the host galaxy. These high-redshift galaxies are considerably more blue than their lower-redshift counterparts, and are considered to be undergoing very active star-formation.

It is worth approaching this apparently general observational result with caution,



The distant radiogalaxy 3C368, whose stellar and radio axes are aligned. These images are taken, *a* in the infrared K band (2.2-μm wavelength), *b*, at radio frequency (5 GHz) and *c*, in the visible (R+V+B bands). (From ref. 8; images *b* and *c* by Djorgovski *et al.*⁹.)

1. Barbad, M. A. *Rev. Biochem.* **56**, 779–828 (1987).
2. Willumsen, B. M., Norris, K., Papageorge, A. G., Hubbert, N. L. & Lowy, D. R. *EMBO J.* **3**, 2581–2585 (1984).
3. Treton, A. M. & Mulshine, J. L. *Nature* **337**, 406 (1989).
4. Hancock, J. F., Magee, A. I., Childs, J. E. & Marshall, C. J. *Cell* **57**, 1167–1177 (1989).
5. Schafer, W. R. *et al. Science* **245**, 379–385 (1989).
6. Casey, P. J., Solski, P. A., Der, C. J. & Buss, J. E. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
7. Anderregg, R. J., Betz, R., Carr, S. A., Crabb, J. W. & Duntze, W. J. *biol. Chem.* **263**, 18236–18240 (1988).
8. Power, S. *et al. Cell* **47**, 413–422 (1986).
9. Gutierrez, L., Magee, A. I., Marshall, C. J. & Hancock, J. F. *EMBO J.* **8**, 1093–1098 (1989).
10. Clarke, S., Vogel, J. P., Deschenes, R. J. & Stock, J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4643–4647 (1988).
11. Magee, T. & Hanley, M. *Nature* **335**, 114–115 (1988).
12. Beck, L., Hosick, T. J. & Sinensky, M. *J. cell. Biol.* **107**, 1307–1316 (1988).
13. Maitese, W. A. & Sheridan, K. M. *J. cell. Physiol.* **125**, 540–558 (1985).
14. Fairbanks, K. P., Barbu, V. D., Witte, L. D., Weinstein, I. B. & Goodman, D. S. *J. cell. Physiol.* **127**, 216–222 (1986).

however. First, as Stockton and Lilly point out, the concept of optical position angle is applied not only to truly elongated continuum sources, but also to objects with discrete but aligned components and even to those which are almost circular (with present resolution). Second, the sample that is the basis of most of this work is the 3C radio catalogue, which includes only very luminous objects and is probably not typical of most sources at high redshifts. Third, we note (ref. 5 and S.G. Neff, personal communication) that an increasing number of radio sources are unresolved with increasing redshift (over 65 per cent at the highest redshifts), so that the alignment conclusions, by definition, do not apply even to all known objects. Finally, in most (perhaps all) low-redshift cases, the host galaxies are in a state of tidal interaction with a companion^{6,7}, which is probably the cause of the nuclear activity, but may also independently trigger widespread star-formation. (This also allows the nonalignment of nuclear and galaxy angular momentum). We do not yet know if there are other or more significant nuclear 'triggers' at high redshift. Despite these caveats, it is clear that the new results offer us significant new insights into the processes and cosmic evolution of activity in galactic nuclei.

Rees, in his new work¹, is concerned with the cases in which alignment is observed. He examines the mechanisms that give star formation as a result of nuclear radio activity. His model requires a two-phase gas, such as would be present as a galaxy grows by the gravitational collapse of primordial gas: hot clouds heated by the collapse, and cold clouds trapped by these. Rees suggests that the shock front of a fast moving radio jet (travelling at a tenth of the speed of light) emerging from the nucleus would compress the cold gas, leading to a well synchronized burst of star formation along the jet's path. The mechanism is plausible and must apply in the general situation described. Certainly it can explain the close correspondence between optical continuum and radio structure in the confines of an infall from perhaps 100 kiloparsec. We do not yet know how common this situation is in the observed systems, some of which, being at low redshift, are being observed as they were long after the main epoch of galaxy formation. Indeed, even the earliest known galaxy (at a redshift of 3.4) shows evidence of an old stellar population. But the observed trend of reduced alignment with increasing galactic age (reduced redshift) is certainly expected from Rees's model, which is most effective in the early stages of a galaxy's lifetime: the hot, blue stars created by the radio jet will last 10^8 years or less, and the inflow of gas will carry less luminous stars into the galactic centre (or elsewhere) over a similar period.

A fuller understanding of the processes and their evolution now also requires a look at the ionizing effects of the central radiation, the radio jet material and the young hot stars themselves. Estimates suggest that all these have a comparable influence, and may dominate in different circumstances or at different stages of a radio outburst. A better statistical picture of the individual evolution of radio sources, and how this changes over cosmic time, is needed. At present, we do not know if all radio sources become extended, or whether the bulk velocities are 3 or 30 per cent the speed of light, and whether the central radiation is beamed, either relativistically or by physical collimation. We also do not know how much of the structure that we see at sub-arcsecond levels in high-redshift objects is due to gravitational lensing: early results suggest this may be significant. A comparison of

ECOLOGY

An inordinate fondness for ants

Robert M. May

HALDANE's best-remembered remark, that God has 'an inordinate fondness for beetles', was elicited by Jowett's question, at high table at Balliol, as to what his studies had revealed about the deity. Not only are there more named species of beetles than of any other order of organisms (somewhere around 300,000), but Erwin's studies of the beetles found in the canopies of tropical trees suggests the number is, as yet, literally 'inordinate': there may be 10 million or more species of beetles^{1,2}. But Haldane's response contains the unspoken assumption that God is more interested in species diversity than in numbers of individuals, for otherwise Jowett might well have been answered with 'an inordinate fondness for ants'. Preliminary analysis by Tobin *et al.* (Harvard University) of the canopy fauna collected by insecticidal fogging from roughly 1,000 m² of the Manu National Park in Peruvian Amazonia shows that ants represent 70 per cent of all arthropod individuals, and around 50 per cent of the animal biomass by dry weight; only about 10 per cent of all individuals are beetles (although, at something like 15,000 named species, ants remain much less diverse).

This piece of chauvinism on behalf of ants is an apt motif for the *First International Symposium on Interactions between Ants and Plants*^{*}, which ranged broadly over the many aspects of ant-plant associations only just beginning to be fully appreciated. As Davidson and Fisher (University of Utah) observed, the relationships between ants and those plants ('myrmecophytes') that provide

the optical continuum structure and ionized gas structure between unresolved and large and small radio sources will also help distinguish between the principal processes proposed. The data are painstaking to assemble, but they will lead eventually to a direct view of the earliest stages of the history of galaxies. □

John Hutchings is at the Dominion Astrophysical Observatory, Victoria V8X 4M6, Canada.

1. Rees, M.J. *Mon. Not. R. astr. Soc.* **239**, 1P-4P (1989).
2. Baum, S.A. & Heckman, T. *Astrophys. J.* **336**, 681-701 (1989).
3. McCarthy, P.J. *et al.* in *Active Galactic Nuclei* (eds Osterbrook, D.E. & Miller, J.S.) 543-544 (Reidel, Dordrecht, 1989).
4. Stockton, A. & Lilly, S.J. in *Active Galactic Nuclei* (eds Osterbrook, D.E. & Miller, J.S.) 541-542 (Reidel, Dordrecht, 1989).
5. Neff, S.G. *et al. Astr. J.* **97**, 1291-1305 (1989).
6. Hutchings, J.B. *Astrophys. J.* **320**, 122-134 (1987).
7. Sanders, D.B. *et al. Astrophys. J.* **325**, 74-91 (1988).
8. Chambers, K.C. *et al. Astrophys. J.* **329**, L75-L79 (1988).
9. Djorgovski, S. *et al. Astr. J.* **93**, 1307 (1987).

them with food and housing, in the form of structures such as extrafloral nectaries and chambers in swellings at the bases of thorns, are still treated in most ecological texts as beautiful cameos, unique and curious examples of symbiotic mutualism largely unrelated to mainstream ecology.

But more detailed studies show patterns and complexity paralleling those in other ecological systems. Thus, Davidson and Fisher have studied the associations of ants with six different species of *Cecropia* plants in Amazonian Peru to test ideas, developed by McKey, about the manner in which habitat-correlated variation in the growth patterns of plants determines their relative investment in chemical and biological defences³. Chemical defences, such as tannins and lignins, are relatively costly for a plant to synthesize, but have relatively low maintenance costs and so represent better investments for plants with long-lived leaves. Conversely, the substances produced as benefits to ants by myrmecophytes, not being toxic, may be reclaimed from senescent leaves and shunted to new growth, so that ant defences may be expected to replace phenological defences in plants with continuous leaf production and rapid leaf turnover.

Broadly, these are the patterns Davidson and Fisher found. They made three pairwise comparisons between *Cecropia* species growing at high light intensity, either in disturbed areas along water-courses or in clearings made by people, and species established mainly in small light gaps in forests. Glossing over many details, it appears that the former species have fast growth, high leaf turnover and

^{*} *First International Symposium on Interactions between Ants and Plants*, Oxford, 6-8 July, 1989.

tend to favour biotic defence by means of ants, whereas the latter grow more slowly, have longer-lasting leaves and tend to invest more in chemical defences. Davidson and Fisher went further to suggest how this broader perspective on ant-protection mutualism can explain patterns of parapatric speciation among *Cecropia*, as plants have colonized new habitats. This theme was also developed by McKey (University of Miami), who studied three species of *Leonardoxa* occurring from Eastern Nigeria to Gabon, showing that the evolutionary history of this system — the phylogenetic relations among the plants and among their associated ants — can be inferred from the character of the vegetative and reproductive parts of the plants, particularly as related to the interactions with different ant species.

Another aspect of the complexity of ant-plant systems is the entrainment of many other species in such associations. The larvae of two groups of butterflies, the lycaenids and the riodinids, form mutualisms with ants. DeVries (Smithsonian Tropical Research Institute) showed that ant-tended riodinid species typically use host plants with extrafloral nectaries, feeding on meristem tissues and drinking the nectar (see figure); riodinid species not tended by ants typically are not found on myrmecophytes. The ant species that tend riodinid caterpillars are harvesters of secretions who also tend homopterans (as well as extrafloral nectaries). All of this points to riodinids as examples of the invasion and exploitation of ant-plant mutualisms.

The mutualism between *Piper* plant species and the ant *Pheidole bicornis* is the only known symbiosis in which the plant is believed to produce food for ants only in their presence, in the form of single-celled food bodies within hollow petiole chambers. But Letourneau (University of California at Santa Cruz) described her recent discovery of a *Phyllobaenus* beetle whose larvae appear to have 'broken the code', inducing the plant to produce food bodies in the absence of ants. These beetles can also disrupt ant colonies, resulting in increased herbivore damage to the plant. *Phyllobaenus* seems to be a parasite of a mutualism.

Cushman, Addicott and Breton (University of Alberta) enlarged on the theme that the dynamics of the interactions between species depend strongly on the ecological setting, focusing on ant-plant-homopteran mutualisms. Their detailed investigations of two particular ant-homopteran associations in the western United States showed the dynamics to depend on the density and age-structure of the populations, to be affected by competition and to depend on quality of host plants. Bristow (Michigan State University) carried this further, to address the paradox of why, since essentially all

An Ecuadorian riodinid caterpillar (*Nymphidium* species) drinking extrafloral nectar whilst being protected against predators by ants, which also deter herbivores from feeding on the plant.

aphids feed on phloem sap and produce honeydew, are so few species of aphids tended by ants — fewer than a quarter of aphid species in the Rocky Mountain region.

Bristow showed that phylogenetic inertia cannot explain the observed patterns, because most subfamilies of aphids contain both tended and untended species. Nor are the patterns easy to explain on the basis of varying selective pressures to avoid predation, because ant bodyguards are only one option among many, often mutually incompatible (as, for instance, living in galls on plants). Bristow developed the tentative idea that the quality of the aphid's host plant may be particularly important, through its effect on the honeydew the aphids produce: tending is much commoner for aphids on some plants (*Melus* and *Prunus* species) than on others (*Betula* and *Acer* species), and a detailed study of an occasionally tended aphid species (*Apis nerii*) feeding on oleander (*Nerium oleander*) showed that aphid colonies feeding on 'high quality' parts of the plant (floral tips rather than leaf meristems) are significantly more likely to attract ants.

Many papers and posters at the meeting dealt with actual or potential applications of ant-plant research. Whittaker (University of Lancaster), for example, surveyed the tradition — long-established in Europe, growing in North America — of introducing wood ants to control insect pests in forests. Emphasizing experimental studies involving broad-leaved trees, he showed that ants of the genus *Formica* do indeed reduce the damage done to trees by herbivorous insects, although the effects vary with season and with the composition of the insect community. Whittaker noted that such manipulative experiments with ants provide a natural and large-scale way of evaluating the impact of herbivorous insect on trees. Dejean (Université Paris-

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Nord) reported related studies of two species of predatory ants important in protecting orchards in the tropics. The impact of the African weaver ant can be dramatic: these creatures live in nests made from leaves woven together with silk spun by the larvae, and Way has recorded colonies consisting of as many as 150 nests containing close to one million individuals⁴.

The opening address by Beattie (Macquarie University) summarized recent advances, but also indicated aspects of ant-plant associations seeming to need more attention. These include the genetics of traits with specific adaptive significance, the biochemical basis of plant-borne rewards for ants, pollination by ants and the extent to which ants affect the gene flow or the spatial organization of gene frequencies in plant populations. Of particular interest to me was Beattie's reminder of the relative neglect of demographic models tailored for ant-plant associations. Andersen (CSIRO Tropical Research Centre, Australia) made a beginning on the task, noting that in their foraging behaviour, patterns of growth and reproduction and in their generally modular nature, ant colonies resemble plants more than animals. This is one more specific instance of the general trend towards constructing population models that go beyond the older Lotka-Volterra metaphors, but which stop short of mind-boggling complexity and a snowstorm of parameters by dealing with specific categories of life histories⁵. □

Robert M. May is a Royal Society Research Professor at Imperial College London and at the Department of Zoology, University of Oxford, Oxford OX1 3PS, UK.

1. Erwin, T.L. *Coleopt. Bull.* **36**, 74–82 (1982).
2. May, R.M. *Science* **241**, 1441–1449 (1988).
3. McKey, D.B. *Biotropica* **16**, 81–89 (1984).
4. Way, M.J. *Bull. Ent. Res.* **45**, 93–112 (1954).
5. Roughgarden, J. in *Perspectives in Ecological Theory* (eds Roughgarden, J., May, R.M. & Levin, S.A.) 203–226 (Princeton University Press, 1989).

Beauty of the standard model

David J. Miller

REPORTS of Z^0 production¹ at the new electron-positron colliders, the linear collider (SLC) at Stanford and the LEP ring at CERN near Geneva, provided an upbeat start to the European Physical Society's recent high-energy physics conference*. Even though the number of Z^0 's recorded so far is only a taste of what is to come, their significance is already apparent in the improved parameter values they indicate for the electroweak part of the 'standard model', which seeks to describe the fundamental physical interactions. Once again, the standard model has survived for another year the tests conducted in several experiments described at the conference, and new ideas emerged on ways to explore and test it further.

The two most pressing experimental questions remain. What are the masses of the top quark and the Higgs scalar particle, if either or both of them exist? There are two well established doublets of quarks — (down, up) and (strange, charm), each with charges $(-1/3e, +2/3e)$, where e is the charge on a proton. This doublet structure is important in understanding how the quarks couple to the W and Z^0 bosons, the quanta of the weak part of the electroweak interaction, which mediate such processes as nuclear β -decay and neutrino scattering. The quarks in the second doublet are heavier than those in the first, and the charm quark is much heavier than its strange partner (which is why it took an additional 20 years to be discovered). The fifth, beauty quark, which is even heavier, has been well established as a $-1/3e$ charged object for some years. Results from the 60-gigaelectron-volt (GeV) Tristan electron-positron collider in Japan (T. Fujii, KEK) confirmed earlier indications (from PEP at Stanford and PETRA at the German accelerator centre, DESY) that the directional asymmetry in beauty-antibeauty production with respect to the initial positron direction has exactly the size and sign expected if the beauty quark is also a member of a doublet — with the undiscovered top quark as its heavier partner with $+2/3e$ charge (see table). (The alternative names for this pair — bottom and truth — which have distracting connotations, have been dropped.)

Direct searches for decays of particles containing top quarks have continued at the proton-antiproton colliders. The 'signature' of such decays would be isolated high-energy electrons or muons associated with jets of strongly interacting particles such as pions in events with large missing transverse momentum. The decay

dynamics can be predicted so the signatures are well defined, but the rate of production falls off as the assumed mass of the top quark is increased. The absence of events has raised the lower limit on the mass to 61 GeV (CERN UA1 experiment), 67 GeV (CERN UA2 experiment) or 77 GeV (Fermilab CDF experiment). The better limit from Fermilab reflects the fact that the collider there gives a higher centre-of-mass energy ($1.8 \text{ TeV} = 1,800 \text{ GeV}$), than CERN's (640 GeV). Better statistics over the next few years should push Fermilab's limit well beyond 100 GeV.

The Higgs scalar field² was introduced to provide the link between the weak forces, carried by the massive W and Z^0 , and the electromagnetic force carried by the massless photon. It also provides a mechanism to account for the masses of many particles, including the everyday electron. Within the standard model (as reviewed by L. Maiani, Univ. Rome and R. Barbieri, Univ. Pisa) theoretical upper limits can be placed on the mass of the free Higgs particle, correlated with similar limits on the mass of the top quark. Several workers (including D. Haidt, DESY) have made fits to the full set of available data, incorporating new neutrino-scattering results from the CHARM2 experiment at CERN and the

ratio of the rates for W and Z^0 production at the proton-antiproton colliders. They calculate that the top mass must lie between approximately 100 and 160 GeV. As measurements of the masses and widths of the W and Z^0 improve, such fits may be able to predict the mass of the Higgs as well, always assuming that the standard model is true.

As explained above, there seem to be at least three 'generations' of quark doublets in the standard model. These are matched with three generations of lepton doublets; (electron, electron neutrino), (muon, muon neutrino) and (tau, tau neutrino). The SLC measurements of the shape of the Z^0 resonance favour the existence of only these three kinds of neutrino¹. This conclusion is supported by the ratio of the W and Z^0 production rates in proton-antiproton collisions (CDF at Fermilab, UA2 at CERN). Some of the uncertainties in interpreting the W/Z^0 ratio were removed by data reported by NMC (New Muon Collaboration) workers at CERN (J. Nassalski, Univ. Warsaw) on the momentum distribution of the quarks inside a proton or an antiproton. (Such distributions are needed when calculating collision rates at a proton-antiproton collider because every quark which takes part in an interaction carries only a fraction of the momentum of the projectile, the rest being shared with the 'spectator' quarks and other material. Electron-positron colliders do not have this problem.) As M. Rees (Univ.

Spruces not warming to the greenhouse



THE environment's response to long-term climate changes is much debated. With the aid of mediaeval spruce remains such as shown above, Serge Payette and colleagues discuss on page 429 of this issue the effect of a series of changes over the past 1,000 years indicating a period of cool weather (AD 1305–1435), a warm period (1435–1570) and the little ice age (1570–1850). From analyses of the tree rings and growth forms of subfossil spruces from northern Canada, the authors find that at these high latitudes (where greenhouse warming should be particularly pronounced), this species flourished even at marginal sites during the warm period preceding the little ice age, whereas the recovery during the present warming is significantly delayed. In part, ecological damage sustained during the past cold 300 years may account for the delay. □

* High Energy and Particle Physics, Madrid, 6–13 September 1989.

Cambridge) pointed out in his comprehensive review of the connections between particle physics and cosmology, confirmation of the existence of only three kinds of light neutrino would give an important extra constraint to the cosmologists' picture of how helium nuclei

of mesons containing beauty quarks were produced. The beauty quark usually decays to the heavy charm quark (third generation to second generation). If it goes to the lighter up quark (third to first generation), there is more energy available for the final particles. Both experiments

find about the same number of extra electrons at the high-energy end of the spectrum, just beyond the limit allowed in decays to charm.

Higher energy means new possibilities in particle physics, so that the news (J. Gunion, Univ. California, Davis) that the American 20 TeV

LEP tunnel to be ready by 1997. But it was interesting that in his talk on new accelerating techniques, W. Schnell (CERN) suggested that a linear electron-positron collider of 200 GeV on 200 GeV might also be a serious possibility: fits to the standard model predict that the top-quark mass is well below 200 GeV, so that such a machine could produce top-antitop pairs and measure the top mass very precisely. This would, in turn, put very tight constraints on the mass of the Higgs particle. The machine might even produce detectable Higgs particles — a result as important to cosmology as to particle physics. The feasibility of such a machine and the scientific case for it are still a little vague, but over the next year or so practical lessons from the Stanford Linear Collider and improved parameters for the Z^0 will make this less so. □

David J. Miller is in the Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK.

1. Miller, D.J. *Nature* **340**, 677–678 (1989).
2. Aitchison, I.J.R. *Phys. World* **2**, 29–34 (1989).
3. Miller, D.J. *Nature* **333**, 799–800 (1988).

Particle type	The particle family			
	quark		lepton	
Charge	$-\frac{1}{3}e$	$+\frac{2}{3}e$	$-e$	0
Generation:				
third	b	t?	τ	$\nu_\tau(?)$
second	s	c	μ	ν_μ
first	d	u	e	ν_e

were formed during the hot Big Bang.

New results from Fermilab seem to have reopened a debate about symmetry-breaking in the standard model. Long-lived neutral kaons are known to decay about one time in a thousand to a final state of two pions. This is a highly suppressed process which violates a simple quantum mechanical rule called CP conservation. The standard model can explain this — as long as there are three generations of quark doublets — and it predicts that the ratio of the decay rates of long-lived and short-lived neutral kaons to two neutral pions will be slightly different from the equivalent ratio for decays to two charged pions. At Madrid, workers on CERN's NA31 experiment confirmed their earlier result¹ which established this difference. But the Fermilab E731 collaboration now claims, with similar precision, to see no difference between the two ratios. The results are only two standard deviations apart, so there is no crisis yet. But it is important to know the mechanism for CP violation in kaon decay, because a related process, at much higher energies during the early milliseconds of the Big Bang, provides the only plausible explanation given so far for the fact that our Universe is made of matter without equal amounts of antimatter.

The standard model explains nonconservation of CP in kaon decay by assuming that virtual beauty quarks make a very small contribution to the decay amplitude. This will happen only if the beauty quark can decay into an up quark, via the weak interaction. The ARGUS group at DESY had claimed two years ago to have seen evidence for such decays, with the rather unlikely production of antiprotons, among other particles, in the final state. That result has now been withdrawn but ARGUS, and the CLEO collaboration at the Cornell electron-positron machine CESR, reported evidence for the much more plausible decay of a beauty quark into an up quark accompanied by an electron and a neutrino. They look at events in which particle-antiparticle pairs

on 20 TeV proton-proton collider, the SSC (Superconducting Supercollider), is to receive its first serious funding from Congress, with the earliest starting date for physics in 1998, was warmly received. J.J. Thresher and K. Potter (CERN) pointed out that Europe is well placed to build an 8 TeV on 8 TeV collider in the

GROUP THEORY

Symmetry breakthrough

Ian Stewart

SYMMETRIES abound in nature and technology. Most living creatures have bilateral symmetry, at least to a good approximation and neglecting internal structure; the same goes for most aircraft. Cylinders, squares, hexagons and other highly symmetric forms occur in innumerable scientific problems. Physicists and chemists in particular have found that the symmetry of a system has a considerable influence on its behaviour. Conventional wisdom, going back to Pierre Curie, holds that symmetric systems behave in a symmetric fashion. It is often useful to bear Curie's principle in mind, but it is now known that for many systems, its naive application may lead to incorrect results. The reason is the occurrence of a phenomenon known as spontaneous symmetry breaking. This behaviour, at first sight paradoxical, occurs when a perfectly symmetric system takes up a state with less symmetry. Recent work of M.J. Field and R.W. Richardson, (*Arch. rational Mech. Anal.* **105**, 61–94; 1989) considerably increases our understanding of the general mechanism by which spontaneous symmetry breaking can occur.

For an example of the phenomenon, imagine a perfect cylindrical shell, compressed from both ends by a uniform force. This system has circular symmetry, that is, it remains unchanged under arbitrary rotations about the axis of the cylinder. However, if the compression is sufficiently large, then both in theory and

in practice the cylinder will buckle into a shape that no longer has complete circular symmetry. Oddly enough, the initial buckling is usually to a state that still has some symmetry, with patterns of ripples around the circumference of the cylinder that repeat at specific angles, like a smoothed version of a regular polygon.

There are many different types of symmetry, and many different examples of spontaneous symmetry breaking. As in the example of the buckling cylinder, there is a marked tendency for at least some, and usually rather a lot, of the initial symmetry to survive. The fundamental theoretical problem is to determine in full generality, for each type of initial symmetry, just how much should survive. The 'folklore' belief, based on the experience of many special cases, is 'as much as possible'. But this is not always correct. Examples of systems with spherical symmetry (P. Chossat *C. r. hebdom. Séanc. Acad. Sci. Paris* **297**, 639–642; 1989; R. Lauterbach *Contemp. Math.* **56**, 217–222; 1986) show that sometimes less symmetry survives than one might expect. The work of Field and Richardson shows that such exceptions are more common than anticipated, and helps to explain why and how exceptions occur.

First, return to the buckling cylinder and ask two questions: where did the symmetry go, and why? 'Why' is largely a matter of stability. In a theoretical model, the circularly symmetric state continues to

exist even at high levels of compression. But it becomes unstable at some threshold value, corresponding to physical buckling. The existence of this threshold is easily verified in specific models. It is to be expected on physical grounds: no realistic cylindrical shell can withstand an arbitrarily high load. If the perfect cylindrical state is unstable, then the system must take up some stable state, which necessarily cannot be circularly symmetric.

'Where' is more subtle. In an idealized model, a whole range of buckled states occur. They all have the same shape, but differ from each other by rotations through arbitrary angles. For example, suppose that in some chosen reference frame there is a stable state in which the cylinder buckles outwards at positions placed 0° , 90° , 180° and 270° around its circumference. Then there is also a stable state for which the buckling occurs at 1° , 91° , 181° and 271° , or for 17° , 107° , 197° and 287° and so on. The original circular symmetry of the single unbuckled state is smeared out across an entire family of buckled states; the individual states lose symmetry but the family as a whole retains it. In practice, tiny imperfections will favour one member of this family rather than another. For example, if the shell has a weakness, this will encourage buckling in that direction. Nevertheless, the idealized model determines the possibilities from which the final selection is made.

To formulate the general problem of spontaneous symmetry breaking, one must make precise the idea of symmetry itself. This is done in the time-honoured way, using group theory. The symmetries of a system are defined to be those transformations that preserve its structure, and the set of all such transformations is called its symmetry group. For example, in a circularly symmetric system the transformations are all possible rotations. The symmetry is said to break if it changes to a smaller group, that is, a subgroup. The fundamental problem of symmetry breaking becomes a question in group theory: given the original symmetry group, which subgroups will occur after the symmetry has broken? The 'folklore' belief becomes the maximal isotropy subgroup conjecture: the subgroups that occur should be the largest ones possible.

Field and Richardson study this question for an important class of symmetry groups, known as Weyl groups. They can be loosely described as the transformations given by a multidimensional kaleidoscope. Weyl groups arise throughout mathematics. In particular, a great deal is known about equations whose symmetries are those of a Weyl group, and Field and Richardson exploit this information to the full. They introduce a new technique, the theory of 'equivariant transversality', which has been worked out over the past decade by Field and E. Bierstone. This lets

them reduce the maximal isotropy subgroup conjecture to a matter of counting geometric dimensions, which they solve.

The end result is that the conjecture is true for some Weyl groups (including those said to be of type A_k , B_k and F_4) but false for others (notably those of type D_k). Indeed they show that its truth for type B_k automatically renders it false for type D_k . The reason is that each group of type D_k is contained as a subgroup in the group of type B_k . For instance, B_4 consists of all permutations of four coordinates, together with all possible combinations of changes of sign: it contains transformations such as $(x_1, x_2, x_3, x_4) \rightarrow (-x_2, x_4, -x_3, -x_1)$. Type D_4 is similar, but contains only half as many transformations, namely those for which the total number of minus signs is even.

Field and Richardson verify that the geometry associated with the group B_4 implies that the maximal isotropy subgroup conjecture is true for B_4 . Therefore, any equation with B_4 symmetry has a known set of solutions, corresponding precisely to the different maximal isotropy subgroups — the ways to retain as much symmetry as possible. Some of these solutions also correspond to maximal isotropy subgroups of D_4 , but others do not. Now any equation with B_4 symmetry can also be considered as having D_4 symmetry (by 'forgetting' the other half of B_4 that does not belong to D_4). So we have found an equation with D_4 symmetry that disobeys the conjecture, as far as D_4 on its own is concerned. A more delicate analysis shows that such solutions persist even if the equation is perturbed slightly, and hence are not just 'accidents' occurring with zero probability.

This example shows that the existence or nonexistence of symmetry-breaking solutions may involve geometric features associated with auxiliary groups, in addition to the group from which or to which symmetry breaks. By using the information gained in these examples, Field and Richardson have constructed many more examples, yet to be published, for which the maximal isotropy subgroup conjecture fails, including some for which it fails in fairly drastic fashion.

The theory of equivariant transversality is a creation of pure mathematicians, invented for reasons of mathematical elegance and completeness, and was not originally motivated by questions in applied science. That it turns out to be precisely the tool required to obtain a basic insight into spontaneous symmetry breaking was not foreseen. So this is an interesting case where a problem that has refused to yield to a head-on attack has succumbed to an indirect approach from a totally unexpected direction. $\square$

Ian Stewart is in the Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK.

Air suspension

AIR TRAVEL these days, says Daedalus, is becoming increasingly slow, tedious and unpleasant. So he proposes a new way of getting through the air. It combines the best features of the balloon and the aerial cableway. His new 'Gasway' uses a hollow, gas-filled, buoyant cable.

The great advantage of this system is that the cable is effectively self-supporting, so the anchoring pylons can be very many kilometres apart. The cable will 'hang' between them in vast, noble, inverted catenaries, rising far into the atmosphere at the centre of each span. To load it evenly, the vehicles moving along it must be quite regularly distributed — a succession of aerial trams rather than one concentrated train.

DREADCO's aerodynamicists are modelling the behaviour of the Gasway under different loads and in various atmospheric conditions. Cunningly, Daedalus will make his trams a trifle tail-heavy, so that they load the supporting cable unevenly. The resulting inclination of the cable will be propagated ahead of the moving tram. It will generate considerable lift on the cable, tending to cancel the weight of the tram. The transient vehicle loads on the cable will thus be minimized, leaving irregular wind forces as the main problem. Modern carbon-fibre-reinforced, plastic-film materials are so strong for their weight that even a long span of cable should safely withstand quite vigorous gusting. It will, however, flex and wander a good deal. But a tram on the cable will in effect experience the wind force average over hundreds of metres of cable, and should be far more stable than a small aircraft or free-floating balloon in the same conditions. And most of the journey will be in the high atmosphere where the wind is pretty steady.

A Gasway will be cheap and easy to operate. Navigation and air-traffic control problems will vanish, and so will huge out-of-town airports: the cable can easily descend straight into a town-centre terminus. And the system will be very secure. A tram cannot collide with other trams on different routes or be hi-jacked to Beirut.

The main problem is the choice of the gas. Such a large volume of helium would be very costly. Daedalus's brilliant solution is to use natural gas, which is mainly methane, much lighter than air, and has to be piped around the country anyway. A combined air-transport and gas-distribution system should be very profitable.

The obvious snag is the risk of fire. Daedalus plans to add a small amount of a halocarbon radical-trap fire-extinguishing agent to his methane, thus making it effectively non-inflammable. This could easily be frozen out again as the gas leaves the Gasway.

David Jones

Rocks that are too hot to handle

SIR—It has recently been announced^{1,2} that the German Science Foundation (DFG) project to drill a bore-hole (KTB) 14-km deep to investigate the deep structure of the Earth's crust has had to modify its objectives, because the unexpectedly high temperatures encountered (about 110 °C at depths of 3,700 m)³ indicate that at the target depth the rock will be too hot to drill. This is despite the fact that the drilling site in the Oberpfalz (see figure) was chosen because of the predicted low temperatures at depth.

Questions arise as to why these predictions⁴, made on the basis of thermal conductivity, magnetic, magnetotelluric and seismic reflection studies, were misleading, and whether the present situation could have been predicted by other means. Shortly before the final choice of site was made, we performed isotope analyses of trace amounts of helium dissolved in deep well waters in the Oberpfalz area⁵. We expected to find ³He/⁴He values close to 10⁻⁸, the average natural production ratio for He in the crust. To our surprise we measured some values greater than 10⁻⁶ (see table). In the absence of excess ³He derived from the decay of tritium generated by atmospheric testing of nuclear weapons, this could be explained only by the presence of a significant proportion of primordial He released from the Earth's mantle. This situation is most commonly found in active volcanic regions, particularly the axial zones of actively spreading mid-ocean ridges, which have a ³He/⁴He ratio of around 10⁻⁵.

These were some of the highest ratios then measured in continental Europe. We took further samples from the Egergraben, a young half-graben structure that displays clear evidence of recent volcanic activity. Here the values were higher still

(see table), very close to those on ocean ridges. The rift is 150 km long and trends ENE–WSW directly into the Oberpfalz.

We concluded⁶ that the only plausible explanation of such He compositions in continental areas was that they had been modified by transient releases of mantle-derived volatiles associated with the emplacement of basaltic magmas in the crust. It is possible that the Egergraben rift is propagating to the west, that we are detecting the precursors to that process and that He serves as a tracer for a geological process that is otherwise undetectable (E.G., R.K.O'N. and E.R.O., manuscript in preparation). Moreover, we showed that if the contemporary emplacement of basalt was as widespread as He measurements made throughout Europe indicated, over geological time such emplacement has probably made a major contribution to the growth of the continents⁷.

The observations of high temperatures in the Oberpfalz seem to provide some support for our hypothesis. They could, however, have a completely different explanation: abnormal crustal concentrations of heat-producing elements or refraction of heat flow by complex geological structures are both conceivable, but neither quantitatively explains the observations. Their contribution should, however, become apparent from further drilling, which should also show whether the temperatures have been raised above conductive equilibrium values by hot fluids ascending convectively through fissures and pore spaces in the rocks.

If the emplacement of magmas was too recent for them to have yet crystallized, they will still be at a temperature of about 1,000 K and the thermal gradient will steepen rapidly as the magmatic body is

approached. If the body has solidified sufficiently, it may be possible to drill into it. In either case the body is probably sufficiently young for the thermal transient associated with its emplacement not to have yet reached the surface and for the helium degassed from it not to have entirely dissipated. The thermal anomaly

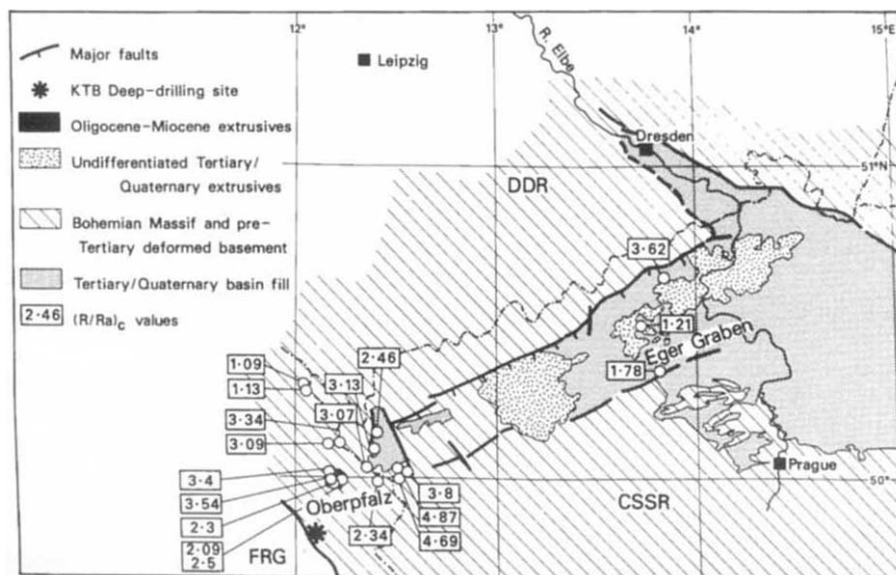
Helium isotope composition of CO₂-rich waters and gases from the Oberpfalz and the Egergraben.

Location	g/w*	He/Ne	(R/R _a) _c †	% mantle helium‡
Bohemian massif				
<i>Oberpfalz, Fichtelgebirge (FRG)</i>				
Kondrau	w	29	3.07	35.5
Thiersheim	w	35	3.09	35.7
Pechofenholz	w	5	2.09	24.2
Sophienreuth	w	47	2.50	28.9
Hohenberg	w	8	3.34	38.6
Pechbrunn	w	49	3.54	40.9
Neualbenreuth	w	366	2.34	27.0
Marxgrun	w	22	1.13	13.1
Wiesau	w	2	2.30	26.6
Alexandersbad	g	84	3.40	39.3
Mahring	w	8	0.66	7.6
Bad Steben	w	184	1.09	12.6
Nagel	w	8	0.66	7.6
<i>Egergraben (CSSR)</i>				
Kynzvalt	g	9	3.80	43.9
Cheb	g	84	3.13	36.2
Smardock	g	3	4.69	54.2
Billina	g	3	1.21	14.0
Soos	g	4	2.46	28.4
Podebrady	g	4	0.90	10.4
Louny	g	481	1.78	20.6
Prameny	g	2	4.87	56.3
Teplice	g	14	3.62	41.8

* Water or gas sample.

† Measured ³He/⁴He ratio in the crust, normalized to that in air (R_a).

‡ Percentage mantle helium calculated according to: mantle He (%) = ((R/R_a)_c × R_a - R_c) × 100/(R_m - R_c), with R_a = 1.4 × 10⁻⁶, R_c (³He/⁴He ratio of radiogenic helium produced in the crust) = 1.5 × 10⁻⁸ and R_m (³He/⁴He in mantle) = 1.2 × 10⁻⁵.



³He/⁴He ratios R (given relative to the ratio in air, R_a) in the Oberpfalz and the Egergraben.

observed could thus be generated by a family of igneous bodies that might be the forerunners of surface volcanoes such as are seen in the Egergraben.

Our prediction, based on the He measurements, that the KTB Project "drilling will provide some surprises"⁸ seems to have been vindicated⁸.

R. K. O'NIONS

E. GRIESSHABER

E. R. OXBURGH

Department of Earth Sciences,
University of Cambridge,
Downing Street,
Cambridge CB2 3EQ, UK

1. Dickman, S. *Nature* **341**, 9 (1989).
2. *The Scientific Driller II* (June) (J. Rühle, California, 1989).
3. Zoth, G. *KTB Report 89-1* 65 (Niedersächsisches Landesamt für Bodenforschung, Hannover, 1989).
4. Weber, K. & A. Vollbrecht *Ergebnisse der Vorerkundungsarbeiten Lokation Oberpfalz, 2-KTB Kolloquium Seeheim/Odenwald 19-21 Sept.* (1986).
5. Griesshaber, E., O'Nions, R.K., Oxburgh, E.R., Lodemann, M. & Paces, T. *Terra Cognita* **7**, 165 (1988).
6. Oxburgh, E.R. & O'Nions, R.K. *Science* **237**, 1583 (1987).
7. O'Nions, R.K. & Oxburgh, E.R. *Earth Planet. Sci. Lett.* **90**, 331-347 (1988).
8. Dept. of Earth Sciences, Contribution No. 1465.

DNA-binding motif

SIR—We wish to point out that a basic motif implicated in sequence-specific DNA binding in the Fos/Jun family of proteins is structurally conserved in the Myc proteins and in other regulatory factors that are proposed¹ to contain amphipathic helix-loop-helix (A-HLH) domains. In addition to the Myc proteins, this group includes the muscle determination factors MyoD1 and myogenin, the kappa enhancer-binding proteins E12 and E47, and the polypeptides encoded by the *Drosophila* genes *daughterless*, *twist* and *achaete-scute* T4/T5.

In the Fos/Jun proteins, the 'leucine zipper' region has been shown to be critical for dimerization, while sequence-specific DNA binding by Fos/Jun requires the integrity of a basic motif immediately amino-terminal to the leucine zipper^{2,3}. Similarly, in the case of the E12/E47 enhancer-binding proteins, integrity of the proposed A-HLH domain has also been shown to be important for protein dimerization and DNA-binding activity¹. The carboxy-terminal region of c-Myc (amino acids 369–439), containing both the leucine zipper and A-HLH motif, has been shown to be sufficient for protein oligomerization *in vitro*⁴.

The incorporation of domain definitions in a visual comparison of the region just amino-terminal of the dimerization motifs in the Fos/Jun and A-HLH proteins

	Basic Motif DNA-binding	Fos/Jun "Leucine Zipper"	
	-----	-----	
c-Jun	ERIKAERKMRNRRIAASKCKRKRKLERIARLEEKVKTLKAQNSELASTANMLREQVAQL...		
			
JunB	ERIKVERKRLNRRLAATCKRKRKLERIARLEEDKVKTLDAENAGLSSAAGLREQVAQL...		
			
JunD	ERIKAERKRLNRRIAASKCKRKRKLERISRLEEKVKTLKSQNTLASTASLREQVAQL...		
			
FosB	EEEKRRVRNRNRNKAALAAKCRNRRELTDRLQAETDQLEEEKAELESEIAELQKEKERL...		
			
Fra-1	EEERRVRNRNRNKAALAAKCRNRRELTDRLQAETDQLEDEKSGLQREIEELQKQKERL...		
			
c-Fos	EEEKRRIRNRNRMAAAKCRNRRELTDRLQAETDQLEDEKSAQLQTEIANLKEKEKL...	Myc A-HLH helix II/"zipper"	
			
c-Myc	PRSSDTEENVKRNTHVLEKRRNELKRSFFALRDQIPLEN-NEKAPKVVILKKATAYILSVQAEQKLISEEDLLKRRRLQKHLKLEQLNSCA		
			
N-Myc	PRNSDSEDSERRRNHNILERQRRNDLRSSFLTLDHVPVLV-NEKAAKVVILKKATEVHSLQAEHQQLLEKEKQLARQQLKKIEHART		
			
L-Myc	PVSSDTEVTNRKNHNFLEKRRNRNDLRSLFALRDQVPTLASCS-KAPKVVILSKALEYLQALVGAEKRMATEKRLCRQQLQKRIAYLSG		
			
MyoD1	ACKRKTITNADRRKAATMRERRRLSKVNEAFETLKRCSTSSNPQ--RLPKVEILRNIRYIEGLQALLR...		
			
E12	PEQKAEREKERRVANNAREKRLVRDINEAFKELGRMCQHLNSEKPKTKLLILHQVSVILNLEQQVR...		
			
E47	LEEKDLDRERRMANNAREKRVVRDINEAFKELGRMCQHLKSDKAQTKLLELQVAVVILGLEQQVR...		
			
da	ERRQANNAREIRIRIDINEALKELGRMCMTLKSDDPKTKLGLINMAVEVIMTLEQQVR...		
			
twist	NQVRMANVRERQRTQSLNDAFKSLQIIPITLP-SDKL-SKIQTLKLATRYIDFLCRMLS...		
	-----	-----	
	Basic Motif	A-helix I	Loop A-helix II
Basic Motif	Consensus: ERRR-----ER-RR-EB		
	BKNK	C	K-DR
	B	B	

Basic motif and oligomerization domain similarities between Fos/Jun and A-HLH gene regulatory proteins. The single-letter amino acid code is used; B represents a hydrophobic amino acid. Identity is noted by dashes, similarities by dots. Amino-acid sequence is shown for mouse c-Jun (254–311), JunB (239–314) and JunD (261–318) (see ref. 6); mouse FosB (154–211), Fra-1 (106–163) and c-Fos (136–195) (see ref. 7); human c-Myc (337–439), N-Myc (384–462), L-Myc (280–363), MyoD (99–164), E12 (327–393), E47 (327–393), *daughterless* (554–613) and *twist* (357–407) (see ref. 1).

reveals a 17-amino-acid block of structural similarity (see figure). The highest similarity (about 45 per cent with conservative substitutions) is between the Fos and Myc proteins. We noted that this region is positioned exactly in register with hydrophobic amino acids critical for the helix-helix interactions that mediate dimerization by the leucine zipper and (proposed for) the

HLH helix 1. Strikingly, the region overlaps that part of the basic motif in Fos which is necessary for specific DNA-binding to the AP1 recognition sequence^{2,4}. The significance of this similarity is founded on the incorporation of biological meaning (that is, definition of domains) into the search. (Homology search algorithms give this similarity a high quality score when the Fos basic motif region is used in database comparisons.)

On the basis of this structural similarity, we propose that the basic motif in the Myc and A-HLH proteins is likely to encode a sequence-specific DNA-binding function. With regard to the Myc proteins, though a specific DNA recognition site is yet to be identified, our proposal contrasts with suggestions that this function is encoded in the A-HLH domain¹, or in a more amino-terminal region (amino acids 285–318) required for nonspecific DNA binding⁵. The structural similarity between A-HLH and Fos/Jun proteins in this domain suggests that these proteins are part of a large 'basic motif' family evolutionarily related via their DNA-binding structures. This family would contain at least three subsets comprising members that contain only leucine zippers, only A-HLH domains or both.

GEORGE C. PRENDERGAST
EDWARD B. ZIFF

Department of Biochemistry,
New York University Medical Center,
New York, New York 10016, USA

Memory T cells

SIR—Haregewoin *et al.*¹ have reported the exciting finding of a human T-cell line expressing a $\gamma\delta$ T-cell receptor (TCR). The T-cell line (CD2⁺, CD3⁺, CD16⁺) responds to PPD (purified protein derivative) and a mycobacterial heat-shock protein of relative molecular mass 65,000 (65K HSP) in a self-restricted fashion. The line was obtained by stimulation of peripheral-blood mononuclear cells with PPD, followed by depletion of CD4, CD8 and $\alpha\beta$ TCR expressing cells. Cells were maintained by repeated restimulation with autologous antigen-presenting cells, antigen and T-cell growth factor.

We have shown that recall-antigen-reactive 'memory' T cells occur with high frequencies in freshly isolated CD45R0⁺/UCHL1⁺, but not CD45RA⁺, populations of human lymphocytes². Haregewoin *et al.* speculate in reference to this work, that their line may be derived from memory T cells, as it is CD45R0⁺/UCHL1⁺ and CD45RA⁺. We would like to point out that after activation *in vitro*, CD45RA⁺ T cells lose CD45RA (refs 3–7) and acquire reactivity with the CD45R0 reagent UCHL1 (refs 6 and 7). The expression of CD45R0 on J1-2 dependent T-cell lines

and clones is, therefore, a regular finding which does not provide any information on the CD45 phenotype of their precursor cells.

Our group has recently demonstrated the expression of CD45R0 by a proportion of freshly isolated $\gamma\delta$ TCR T cells. This proportion is low in very young children and rises with age⁸, as is the case for $\alpha\beta$ TCR T cells. This observation is consistent with the accumulation of $\gamma\delta$ memory T cells, a possibility that could be tested now that some antigens to which human $\gamma\delta$ T cells respond have been defined^{1,9,10}.

MATTHIAS MERKENSCHLAGER
PETER C. L. BEVERLEY

ICRF Human Tumour Immunology Group,
91 Riding House Street,
London W1P 8BT, UK

1. Haregewoin, A. *et al.* *Nature* **340**, 309–312 (1989).
2. Merkschlager, M. *et al.* *Eur. J. Immunol.* **18**, 1653–1661.
3. Tedder, T.F., Clement, L.T. & Cooper, M. *J. Immunol.* **134**, 2983–2993 (1985).
4. Moore, K. & Nesbitt, A.M. *Immunology* **61**, 159–165 (1987).
5. Serra, H.M. *et al.* *J. Immunol.* **140**, 1435–1441 (1988).
6. Terry, L., Pickford, A. & Beverley, P.C.L. in *Leucocyte Typing 3* (eds McMichael, A. *et al.*) 225–227 (Oxford University Press, 1987).
7. Akbar, A. *et al.* *J. Immunol.* **140**, 2171–2178 (1988).
8. Hayward, A.R., Lee, J. & Beverley, P.C.L. *Eur. J. Immunol.* **19**, 771–773 (1989).
9. Holoshitz, J., Koning, F., Coligan, J.E., De Bruyn, J. & Strober, S. *Nature* **339**, 226–229 (1989).
10. Modlin, R.L. *et al.* *Nature* **339**, 544–548 (1989).

1. Murre, C. *et al.* *Cell* **56**, 777–783 (1989).
2. Kouzarides, T. & Ziff, E. *Nature* **336**, 646–651 (1988).
3. Neuberg, M., Schuermann, M., Hunter, J.B. & Müller, R. *Nature* **338**, 589–590 (1989).
4. Dang, C.V. *et al.* *Nature* **337**, 664–666 (1989).
5. Dang, C.V. *et al.* *Molec. cell. Biol.* **9**, 2477–2486 (1989).
6. Hirai, S.-I. *et al.* *EMBO J.* **8**, 1433–1439 (1989).
7. Zerial, M. *et al.* *EMBO J.* **8**, 805–813 (1989).

People are different

Christopher Longuet-Higgins

The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics. By Roger Penrose. Oxford University Press: 1989. Pp.466. £20. To be published in the United States in November, \$24.95.

ONE cannot imagine a more revealing self-portrait than this enchanting, tantalizing book. The Rouse Ball Professor of Mathematics at the University of Oxford might be thought of as a dusty old don with a formidable reputation in some unheard-of field and an occupational inability to talk about any topic of general interest. Roger Penrose reveals himself as an eloquent protagonist, not only of the wonders of mathematics, but also of the uniqueness of people, whom he regards as mysterious, almost miraculous beings able to burst the bounds of mathematical logic and peep into the platonic world of absolute truths and universal objects.

The book's title is a storm warning for a furious attack on strong AI (artificial intelligence) — the philosophical view that computers 'really' have minds, and that any computational device, even a thermostat, has mental qualities of some sort. From this postulate it would seem to follow that the human mind could well be superseded by a sufficiently powerful computer, a prospect that most people find either alarming or absurd. Penrose decides to demonstrate its absurdity by mathematical argument, the most powerful intellectual weapon known to man. He refers with approval to John Searle's attack on strong AI, but takes an even more extreme view in dissenting from Searle's surprising concession that "Of course the brain is a digital computer. Since everything is a digital computer, brains are too". "It is my intention in this book", Penrose explains, "to try to show why, and perhaps how, this need not be the case".

For his critique of the contention that the human brain is a digital computer Penrose marshalls a range of arguments from mathematics, physics and metamathematics. One of the book's outstanding virtues is the trouble its author takes to acquaint his readers with all the facts they need in order to understand the crucial problems, as he sees them, and all the steps in any argument that underpins an important theoretical conclusion. He

quotes, in his foreword, the dictum that each mathematical formula in any book will cut down its general readership by half, but defies this maxim by quoting a number of equations far greater than the binary logarithm of the world's population. His serious readers will be grateful to him for explaining *why* there are problems in identifying truth with what can be

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Roger Penrose — giving a piece of his mind.

proved, or in equating reality with what can be measured; others will be happy in the company of a guide whose quips they can enjoy even if they cannot always follow his allusions.

In his first attack on conventional AI wisdom Penrose develops an argument very similar to the one that John Lucas first propounded in the early 1960s, to the great annoyance of the computer scientists of that generation. The human mind cannot, he maintains, be a mere computer following through the steps in an complicated algorithm, because there can be no single algorithm for deciding *all* the questions that a mathematician might pose — and answer. In order to explain why, he gives a brilliant and intellectually strenuous account of algorithms and their implementation by Turing machines. But having made his point Penrose unexpectedly expresses some qualms about the relevance of these particular concepts to the philosophy of mind:

It has been helpful in the above description to

use the anthropomorphic term 'know' in reference to an algorithm. However, is it not *we* who are doing the 'knowing' while the algorithm just follows the rules we have told it to follow? Or are we ourselves merely following rules that we have been programmed to follow from the construction of our brains and from our environment? The issue is not really simply one of algorithms, but also a question of how one judges what is true and what is not true. These are central issues that we shall have to return to later.

The next stop on this extraordinary cosmic tour is the kingdom of pure mathematics, illustrated (how better?) by the astonishing Mandelbrot set, and pedagogically enriched by an exposition of complex numbers. The idea of non-recursive mathematics, beyond the reach of feasible computation, is used to illus-

trate the feebleness of proof and the power of insight in the mathematician's quest for truth. The author quite rightly yields to the temptation to reminisce about some of his own discoveries, including the famous 'Penrose tiles', with which one can tile the plane, but only in a non-periodic fashion. He returns to the tiles much later in a visionary passage on how the recently discovered 'quasi-crystals', which give the appearance of possessing five-fold symmetry axes defying the principles of crystallography, might emerge from their mother-liquor by a quantum-mechanical mechanism possibly bearing some analogy to the

way in which the mind distils coherent ideas out of a mass of inarticulate intuitions.

Penrose explains that his book was inspired by the conviction that we shall not understand how our brains work until we have a much better understanding of *physics*. The greater part of the book is, accordingly, devoted to the findings of classical mechanics, relativity and quantum mechanics rather than the more obviously relevant disciplines of philosophy, psychology and physiology. One after another the SUPERB theories of physics, as Penrose dubs them (other theories being merely USEFUL or TENTATIVE) have run into acute difficulties of internal consistency or physical interpretation. His account of the problems and paradoxes of contemporary physics is delivered with all the assurance of a leading cosmologist — one who believes, incidentally, that a revolution in quantum gravity is about to break upon the scientific world. Not only, according to Penrose, could this fill an awkward gap in the quantum theory of observation

Anita Corbin/John O'Grady © The British Council

(relating to the circumstances under which wave functions are deemed to collapse), but it might very well resolve the apparent discrepancy between the time-reversibility of the laws of dynamics and the manifest irreversibility of the second law of thermodynamics and the subjective arrow of time. Central to the latter problem is Penrose's 'Weyl Curvature Hypothesis', which distinguishes sharply between initial and final singularities in cosmic space-time. As for the collapse of wave functions, he speculates that this happens whenever a "significant" amount of space-time curvature causes quantum superposition to fail. The latter is perhaps the most TENTATIVE suggestion in the book, and the author is duly cautious in advancing it.

Having journeyed to the ends of the Universe in search of possible enlightenment, Penrose wistfully bids farewell to big bangs and black holes, and resumes his reflections on the mind — or rather, on the brain as seen through the eyes of his Oxford colleagues in psychology and neurophysiology:

The picture of a superb computing device seems to be presented to us. The supporters of strong AI would hold that here we have a supreme example of an algorithmic computer . . .

Will our hero succumb to this persuasive image? Not he. There is an ace up his sleeve — "the phenomenon of consciousness". But the philosophical problems surrounding this concept prove so bewildering that Penrose abandons logical argument for rumination upon the nature of free will, the anthropic principle, creative inspiration and the sensing of necessary truths. There is just one cat he refrains from putting among the neuroscientific pigeons, namely the suggestion that the brain is actually a 'quantum computer' of the kind recently defined by David Deutsch. He raises this possibility, but concludes that the brain is too "hot" an object to preserve quantum coherence for any length of time.

The Emperor's New Mind might have been merely an attack on a particular group of philosophers or computer scientists, or on some of the wilder claims of the artificial intelligentsia. It might, with more profit, have examined in detail some of the honest hard work that has been done, by theoretical linguists and cognitive psychologists, in attempting to describe as accurately, as elegantly and in as much detail as possible how the human mind actually *does* work; Penrose must know that it is the details in which the soul of a work of art largely resides. As it is, his book is the testament of a brilliant man wrestling desperately — and unashamedly — with the deepest problems of metaphysics. He does not solve any of them, but his questions are greatly preferable to

most people's answers. His book aptly illustrates his own thesis that philosophically inept scientists may have at least as much to contribute to human thought as scientifically ignorant philosophers. □

Christopher Longuet-Higgins is in the Centre for Research on Perception and Cognition, University of Sussex, Brighton BN1 9QG, UK.

In the dock

Robert Temple

The Galileo Affair: A Documentary History. By Maurice A. Finocchiaro. *University of California Press: 1989. Pp.382. Hbk \$50, £31.25; pbk \$12.95, £8.*

Sidereus Nuncius or the Sidereal Messenger. By Galileo Galilei. Translated with introduction, conclusion and notes by Albert Van Helden. *University of Chicago Press: 1989. Pp.127. Hbk \$29.95, £19.25; pbk \$7.95, £6.25.*

THE trial and condemnation of Galileo by the Roman Catholic Inquisition in 1633 is the classic example of the triumph of mediocrity over genius, and the suppression of the truth in the interests of the Great Lie. The central issue was quite clear. In Galileo's own words, it was this: "who wants the human mind put to death?"

Much has been written about Galileo's trial. But it is only now that we can steep ourselves in the atmosphere of the incident and read the very words of the main protagonists, often from their own secret correspondence or from private Inquisition papers. For the first time, all the relevant documents are made available in English in Maurice Finocchiaro's marvellous translation.

There is a pernicious trend amongst historians today which discourages the publication of original source material. Only tedious and opinionated 'discussion and analysis' by self-styled 'experts' is held to be interesting or worthwhile. The reason, of course, is that many historians want to set themselves up as the clergy who will mediate between historical truth and the public. So strong is this trend that some of those who have dared to publish original documents are ostracized and attacked in the way that Cranmer was savaged in the sixteenth century for daring to suggest that ordinary people should be allowed to read the Bible for themselves. Finocchiaro has braved fashion to enable us to step inside the Inquisition proceedings. For this we cannot be too grateful.

Reading through the actual documents for the 20-year period of 'the Galileo Affair', one is struck by the honesty and genius of Galileo and we can see clearly that he could not possibly have done other

than recant. The Medicis of Florence were so deeply in political jeopardy over their passionate support of him against the Pope that if he had been awkward, and made himself a martyr, it would have brought incalculable disaster upon his patrons.

As for the hypocritical and mindless attitude of the Catholic thought police, whose only interest was the relentless imposition of uniformity and destruction of any trace of independent thought, we have the Pope himself as an example. In his secret diplomatic correspondence, the Florentine Ambassador describes this encounter:

I humbly begged His Holiness to agree to give Galileo the opportunity to justify himself. Then His Holiness answered that in these matters of the Holy Office the procedure was simply to arrive at a censure and then call the defendant to recant. I replied: Does it thus not seem to Your Holiness that Galileo should know in advance the difficulties and the objections or the censures which are being raised against his work. . . ? He answered violently: I say to Your Lordship that the Holy Office does not do these things and does not proceed in this way. . . the Holy Office is not in the habit of hearing defenses. . . .

As we all know, the grey men won and Galileo's *Dialogue on the Two Chief World Systems* was not removed from the Catholic Index of Prohibited Books until 1835. Its principles of putting scientific truth over the words of Scripture was not accepted until 1893, and it was not until 1979 that the Church, in a speech by Pope John Paul II, acknowledged that it had erred in condemning Galileo.

For the proper appreciation of Galileo's genius there is no substitute for reading his *Sidereus Messenger*, a short work charged with the electric excitement of discovery. Published in 1610, it is essentially a scientific newsletter announcing fundamental astronomical discoveries that had been made by means of telescopes of unusually high power for the time. Among the revelations were that the Moon's surface is covered in mountains and valleys rather than being smooth, and the existence of the four principal satellites of Jupiter. The sense of euphoria and joy at discovering such things for the first time in mankind's history is fresh and tangible.

The work has never before been made available in its entirety in a continuous form, with full notes and comment. The introduction, translation and notes by Van Helden are a splendid example of the best scholarship and fullest accessibility. With this *Messenger* and Finocchiaro's magnificent book, we can now truly get to grips with the phenomenon of Galileo and what his life and work should mean to us today. □

Robert Temple can be contacted c/o David Higham Associates, 5-8 Lower John Street, Golden Square, London W1R 4HA, UK.

The chocolate box problem

Peter Bryant

Autism: Explaining the Enigma. By Uta Frith. Basil Blackwell: 1989. Pp.204. Hbk £25, \$19.95; pbk £8.95.

THE number of autistic children is tiny (about 4 in 10,000). Yet their plight has always attracted a great deal of attention because the nature of their difficulties poses an intriguing puzzle.

Autistic children have severe social and emotional problems, and they have some striking intellectual difficulties as well. This combination of problems is unusual in itself, but it seems even more curious when one considers the unevenness of their intellectual abilities. They tend to fail dismally in some intellectual tasks, and yet have little difficulty in others. Some display remarkable talents — there are autistic children who draw astonishingly accurate and beautiful pictures, and others who have remarkable musical abilities, and yet can barely speak.

The general interest in autism is reflected in the large amount of research that has been done on the subject, but there has never been an accessible and comprehensive account of this work. Uta Frith, who is one of the leading researchers on childhood autism, has now filled this gap. Her exciting and attractively written book can be divided into two parts. In the first she lays out the facts — the odd combination of symptoms, the possible influence of heredity, the social class and gender differences — and thus sets the puzzle. The aim of the rest of the book is to provide the solution.

In fact two solutions are offered, or at any rate two hypotheses. The first is that autistic children are insensitive to pattern. Most children find it easier to remember a meaningful sentence than a list of disconnected words, but that is not true of autistic children. They are not helped by meaning. They also often fail to spot repetitive patterns in perceptual displays, though such patterns are perfectly obvious to other children. Sometimes the attention that autistic children pay to parts at the expense of the whole actually helps them. They are better than other children at spotting an abstract shape which is hidden within a larger meaningful figure. So Uta Frith argues that their world is fragmented and that it lacks meaning.

Her second theme is that autistic children find it particularly hard to work out what other people are thinking and feeling. She and her colleagues have shown, to take but one example, that autistic children manage perfectly well when asked to arrange some pictures into a

sequence that tells a story about a balloon flying away and bursting on a tree, or about someone going to a shop and buying sweets. But they are at a loss when the story concerns a boy's disappointment at not finding something which he had put in a box for safe keeping (and which someone else had stolen). When other people's expectations and emotions are different from the autistic child's they are beyond his comprehension.

Uta Frith links this difficulty to another peculiarity of autism — the fact that autistic children tend not to indulge in 'pretend play'. Normal children, when playing, quite often decide, for example, that a banana can play the part of a telephone, but as far as the autistic child is concerned it stubbornly remains a banana. Uta Frith argues that one has to know about pretence to understand that two people can have different beliefs about the same event. If I pretend that an empty chocolate box is full, I know that I may believe one thing and the rest of the world another about the box's contents. According to Uta Frith autistic children do not understand pretence and thus cannot grasp the fact that different people have different feelings and beliefs. As a result many of the most important aspects of social communication elude them completely.

This in turn accounts for their social and also for their linguistic difficulties.

Uta Frith's presentation of these two hypotheses is lucid, entertaining and extremely convincing. There is no doubt that the empirical evidence which supports her ideas is the most exciting that there is on the subject of autism. There are a few objections to be made. Autistic children may have a better understanding of pretence than she suggests: some recent work shows that they do take to pretend play if they are prompted to do so. There is also a certain tension between the two hypotheses. For example autistic children easily rearranged pictures into a story about a balloon bursting, and yet according to Uta Frith's ideas about fragmentation it should have been quite difficult for them to do so. Her explanation is that they have difficulty with broad but not with narrow contexts: but how does one decide what is broad?

These are minor quibbles while the book is without doubt a major achievement. It will have an immense influence, and all to the good, on future research on this topic. □

Peter Bryant is in the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.

Signal survey

Gunnar von Heijne

Protein Targeting. By Anthony P. Pugsley. Academic: 1989. Pp.279. \$45, £28.50.

PROTEIN targeting — that is, the means by which newly made proteins are directed to their final destinations inside or outside the cell — has been among the hottest subjects in molecular and cell biology for more than a decade. Yet it is only with Anthony Pugsley's book that a well-written, integrated overview of the field has become available.

Pugsley starts with two introductory chapters on protein traffic in eukaryotic and bacterial cells, and the main experimental techniques used to study these processes. He then takes a deep breath and dives into his subject, following the sometimes tortuous paths of proteins from the cytoplasm through all the intricacies of the secretory pathway or into the depths of nuclei, mitochondria and chloroplasts.

Throughout, the emphasis is on the structure of the routing signals and on how these are decoded by a battery of cytoplasmic, membrane-bound and intra-organellar proteins that ensure the specificity and vectoriality of targeting and transport. Pugsley provides a unified picture, based on the latest results, with cytoplasmic factors stabilizing 'transloca-

tion-competent' conformations of the newly synthesized proteins, organelle-bound receptors recognizing their cognate routing signals, integral membrane proteins possibly serving as translocation machines, and 'chaperones' inside the organelles assisting in the final steps of folding and assembly.

Also discussed are the vesicular communication links between the different stages of the secretory pathway — endoplasmic reticulum to Golgi to trans-Golgi network to plasma membrane or lysosome — and the complementary processes of endocytotic uptake of proteins and other molecules from the cells' surroundings.

The text ends with an appendix describing various ways that our knowledge about protein targeting can be put to practical ends: secretion of heterologous proteins by bacterial, fungal and animal cell cultures; herbicide-resistance by manipulation of chloroplast-targeted proteins; and specific targeting of toxins into the endocytic pathway of tumor cells.

As a befitting finale there is an impressive bibliography with some 1,300 references, essentially complete up to and including 1988. Let us hope that the royalties Pugsley can expect from his book will go at least some way towards covering his no doubt exorbitant photocopying bills. □

Gunnar von Heijne is in the Department of Molecular Biology, Karolinska Institute Center for Biotechnology, Novum, S-141 52 Huddinge, Sweden.

Talk across the barriers

Andrew Cossins

Membrane Transport: People and Ideas.

Edited by Daniel C. Tosteson. *American Physiological Society*, 9,650 Rockville Pike, Bethesda, Maryland 20814, 1989. Pp.414. \$75. Distributed by Oxford University Press in Britain, £50.

IT MAY be hard for the membrane biologists of today to appreciate fully just how far and how rapidly their subject has progressed in this century. In the 1920s the very notion of a limiting membrane around cells was still controversial. It was only 57 years ago that the bilayer was firmly recognized as the basic ultrastructural component of biomembranes, and only 17 years have passed since the inclusion of protein in the bilayer was formally conceptualized by Singer and Nicholson. Since that time an increasing number of transport systems have been dissected and characterized both at the molecular and submolecular level.

The origins and development of the principles that underpin modern studies are given extensive treatment in *Membrane Transport*, a multi-authored volume written in celebration of the Centenary of the American Physiological Society. The book is wide ranging, starting with detailed accounts of the origins of the bilayer/unit membrane concept, moving through the principal modes of transport across membranes to ion selectivity and bioelectricity. Finally, the emergence of concepts in whole-cell phenomena such as cell-cell communication, epithelial transport of fluid and electrolytes are described.

The authors form a distinguished cast. Among them are people who figured prominently in the early development of membrane biology between 1930 and 1960, as well as others who are known for their work of the past 20 years. This breadth of perspective is quite unusual in contemporary books. Davson's article extends back into the middle of the 1800s to describe the first experimental evidence for the existence of membranes in plant cells, but then moves through to the mid-twentieth century and ideas about facilitated diffusion of ions across membranes.

Perhaps as a means of breathing life into an otherwise factual and rather dry history, the contributors were invited to give an anecdotal account of their roles in the scientific developments, including an impression of the personalities who influenced their work. It is this that gives the book its unusual flavour, although not all of the authors have used the device to



Sitting pretty — Isaac Newton, "a very great head of school but not pompous", depicted in a Japanese print c. 1869. The illustration comes from the new paperback edition of *Let Newton Be!* (Oxford University Press). The book was reviewed in *Nature* **337**, 28 (1989).

its full effect. Lowenstein's charming account of his work on cell-cell communication is a model of the genre, full of anecdote and personal observation, and giving a real feel of the excitement that follows a chance but crucial observation. Similarly, the article by Solomon on the practical difficulties of pre-war science is not only easy to read but gives a good impression of the 'heroic' nature of the research before the industrial production of scientific instruments got underway after the Second World War (his first task was to build a cyclotron for the production of radioisotopes!). Robertson's highly personal account of his work in the early days of the electron microscope and the unit-membrane concept is also enjoyable, especially as it relates to more recent concepts of ultrastructure. It is, though, twice the length of any other contribution and as such is tiresomely long.

Some of the other articles are disappointingly bland, giving more straightforward treatments of the progression of ideas with little personal embellishment save recording of the appearance or disappearance of colleagues as the story progresses. As a result they resemble rather too closely the review articles found

in contemporary multi-authored series.

Membrane Transport will provide good reading for postgraduate students and researchers who like to place the names that they frequently meet in the research literature into historical context. But the book also provides much that is of more than passing interest to those who make a career out of science. For example, there are numerous snippets of information and advice about working and publishing in a competitive scientific environment, advice which younger scientists should read with care. In addition, the recognition by more than one author that chance encounters and observations have played a large part in their achievements is cheering. But the most sobering notion is that provided by Danielli (quoted in Davson) in respect of the inevitability of scientific discovery. Having observed that the lipid bilayer was independently discovered at least three times, he concludes that "No scientist can afford to be arrogant about the degree of originality he achieves". Amen. □

Andrew Cossins is in the Department of Environmental and Evolutionary Biology, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK.

How the ear's works work

A. J. Hudspeth

The senses of hearing and equilibrium depend on sensory receptors called hair cells which can detect motions of atomic dimensions and respond more than 100,000 times a second. Biophysical studies suggest that mechanical forces control the opening and closing of transduction channels by acting through elastic components in each hair cell's mechanoreceptive hair bundle. Other ion channels, as well as the mechanical and hydrodynamic properties of hair bundles, tune individual hair cells to particular frequencies of stimulation.

THE internal ear is an evolutionary triumph of miniaturization, a three-dimensional inertial-guidance system and an acoustical amplifier and frequency analyser compacted into the volume of a child's marble. That we are ordinarily oblivious to what goes on within the labyrinth reflects the ear's reliable performance: our bipedal stance, our capacity for locomotion on foot and in vehicles and our ability to communicate vocally, all stem from the ear's incessant activity. The internal ear, in turn, owes its success to its complement of hair cells, the epithelial receptors responsible in the cochlea for our sensitivity to sound, in the utricle and saccule for our perception of linear accelerations, and in the three semicircular canals for our appreciation of rotatory accelerations.

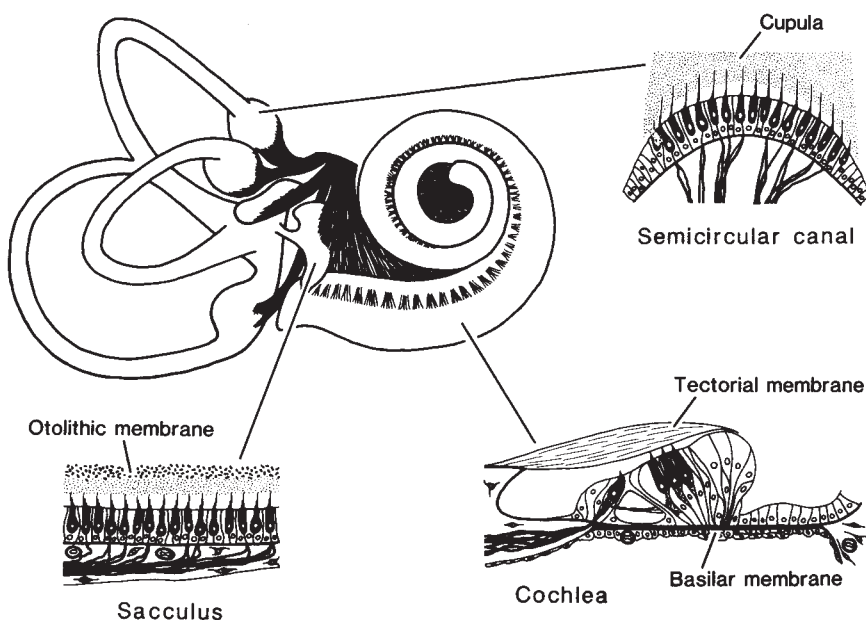
The evolutionary success of hair cells is advertised by the persistence of structurally similar internal ears in all vertebrates. The volume of the human nervous system devoted to the auditory and vestibular systems also testifies to the ear's importance. The 32,000 hair cells of the two cochleae¹ and the total of 134,000 cells found bilaterally in the organs of equilibrium² forward their signals to the large cochlear and vestibular nuclear complexes of the brainstem's medulla and pons. The cochlear nuclei initiate a complex network of rostral projections through the

pons, midbrain and thalamus to the cerebral cortex; the vestibular nuclei send information particularly to the cerebellum, oculomotor system and spinal cord.

The paucity of hair cells also accounts for the exceptional vulnerability of the auditory system. The loss of 166,000 contiguous photoreceptors in a retina would produce a scotoma smaller than the moon's image (and smaller than the blind spot that is already insensibly present). But destruction of 166,000 hair cells would entail disequilibrium and profound deafness, both major impairments. Society is awakening only slowly to the cost of acoustic trauma in both industrial and recreational settings; within the next decade or two, the 'walkman generation' will find itself unexpectedly and ardently interested in auditory pathophysiology.

Hair cells have been the subject of extensive experimentation, particularly during the past 15 years. Techniques have been developed for removing receptor organs from the ears of experimental animals and for maintaining them under microscopic observation in media of defined composition. By mechanically stimulating individual hair cells and recording the resulting electrical responses, researchers have begun to explore the cellular basis of the operation of the ear^{3,4}. Here I shall review the

FIG. 1 Schematic diagram of the right human internal ear from an anterolateral perspective (after ref. 1). Lodged in a cavity of the temporal bone, this membranous labyrinth is immersed in perilymph, which has an ionic composition resembling that of other extracellular fluids. The inner aspect of the labyrinth is lined by a continuous epithelium that secretes and confines the Na^+ - and Ca^{2+} -poor, K^+ -rich endolymph which bathes the receptive surfaces of hair cells. Myelinated nerve fibres, which carry afferent information from the internal ear into the brainstem and efferent information in the opposite direction, converge to form the eighth cranial nerve. Hair cells (darkened for emphasis) occur in different arrays characteristic of the three types of receptor organs. The three semicircular canals are sensitive to angular accelerations about mutually perpendicular axes. Each consists of a toroidal, endolymph-filled cavity interrupted by a septum, the gelatinous cupula. Angular acceleration causes motion of the endolymph with respect to a canal's walls; the displaced fluid deflects the cupula and thus the hair bundles inserted into it. In the cochlea, four rows of hair cells spiral along the elastic basilar membrane in the narrow, 35-mm-long⁷⁸ organ of Corti. Airborne sound sets the eardrum into motion, which is conveyed to the cochlea by the bones of the middle ear. The piston-like motion of the last of these bones displaces the contents of the cochlea's three, fluid-filled internal compartments, thereby flexing the basilar membrane up and down. Hair bundles in the organ of Corti are then stimulated by the shearing motion between the apical hair-cell surfaces and the overlying, gelatinous tectorial membrane. The otolithic



organs, the saccule and adjacent utricle, are sensitive to linear acceleration in vertical and horizontal planes, respectively. The hair cells of each organ occur in a small plaque, or macula; their hair bundles are attached to a gelatinous otolithic membrane upon which lie numerous calcareous otoconia. An acceleration, such as that due to gravity, deflects the dense otoconial mass, then the otolithic membrane and finally the hair bundles.

elegant and subtle ways in which these cells contribute to the functions of the ear, and discuss those of the hair cell's activities that we do not yet understand.

Structure of the hair cell

A human internal ear comprises six separate receptor organs, each sensitive to a particular modality and orientation of stimulation (Fig. 1). Despite this complexity, the hair cells that serve as receptors are of similar form and function in every organ. Each hair cell acts as a biological strain gauge—mechanical stimulation opens ion channels in the cell membrane. The consequent influx of current through these channels alters the membrane potential, which in turn affects the rate of release from the hair cell of a synaptic transmitter. Excited by this chemical, an afferent nerve fibre contacting the basolateral surface of the hair cell transmits to the brain a pattern of action potentials that encodes such features of the stimulus as its intensity, time course and frequency.

From the flattened apical surface of a hair cell projects its hair bundle (Fig. 2), the antenna for reception of mechanical stimuli. This organelle is a cluster of 20–300 cylindrical processes, the stereocilia, each consisting of an actin cytoskeleton^{5–7} ensheathed by a tube of plasma membrane. The stereocilia are hexagonally packed and vary continuously in length across the apical surface of a hair cell; the increment in height between successive ranks of stereocilia is generally constant. A bundle is, therefore, a mirror-symmetrical, bevelled structure shaped like the tip of a hypodermic needle. A single, axonemal cilium, the kinocilium, occurs at the tall extreme of the bundle. The role of this process is uncertain, for it degenerates in the hair bundles of cells sensitive to high frequencies, and transduction persists after its microsurgical ablation⁸.

Within the cytoskeleton of a stereocilium, adjacent actin filaments are extensively cross-linked by fimbrin^{9,10}, which holds contiguous filaments in longitudinal register and separates them laterally by 10 nm with either exact hexagonal symmetry or a less constrained, liquid-like order^{6,7}. Cross-linking renders a stereocilium much more rigid than would be expected for a bundle of actin filaments free to move with respect to one another^{11–13}.

Isodiametric from its distal tip, along most of its length, an individual stereocilium tapers for the micrometre or so above its basal insertion. As the stereociliary diameter diminishes from 100–900 nm (refs 14, 15) to 100 nm or less, the actin filaments decrease from as many as 3,000 to about 20 (ref. 7). These residual microfilaments form a rootlet that inserts into the cuticular plate, a thick mesh of interlinked actin filaments lying beneath the apical membrane surface. If a mechanical force is applied near the tip of a stereocilium and at a right angle to its long axis, the process consequently does not flex, but pivots at its attenuated base¹⁶. When the hair bundle as a whole is deflected by such a stimulus, the constituent stereocilia move in unison,

pivoting at their bases and sliding along one another's shafts. Consonant with this mode of motion, the stiffness of a hair bundle (about 1 mN m^{-1}) is roughly that expected from flexion of the actin cables at the bases of the several stereocilia^{12,13,17}. The stereocilia are not wholly independent of one another, however, for they are interconnected by filamentous linkages, which may account for a bundle being somewhat stiffer along its axis of mirror symmetry than at a right angle to that axis¹⁸.

The hair bundle is among the most elaborate and precisely specified structures in any cell: at a specific position along the basilar membrane, and at a particular location across this membrane, each hair cell has a constant number of stereocilia of fixed length and diameter, disposed in an invariant array¹⁹. Electron microscope observations indicate that this precision results from a morphogenetic program of temporally separate steps²⁰. Rudimentary protrusions initially stud the apical surface of a developing hair cell; beginning with those near the kinocilium, successive ranks of processes then begin to lengthen at increasing intervals. Because all the nascent stereocilia grow at similar rates after elongation commences, the delays in the initiation of growth of successive ranks of stereocilia establish the bundle's bevelled top aspect. While longitudinal growth pauses, individual stereocilia next increase in girth by the circumferential addition of actin filaments about those initially present²¹. After they have achieved their final diameters, the stereocilia resume elongation; the subsequent arrest of growth then fixes a given bundle's definitive proportions. These dimensions vary¹⁹, for bundles cease elongation in a wave that runs up the length of the basilar papilla, leaving hair bundles at the base of the organ much shorter than those at the apex. At least in some internal ears, the height of a hair bundle is a major determinant of characteristic frequency, the stimulus frequency at which a given hair cell is most sensitive^{22,23}. As a result of the sequential morphogenetic events in the internal ear, a temporal gradient in hair-bundle growth, translated into a gradient in bundle height, ultimately establishes a gradient in frequency selectivity.

Our present understanding of the structure and development of the hair bundle raises many fascinating questions. How, for example, are the dimensions of bundles specified? Despite their varied dimensions, each of the hair bundles in the chick's basilar papilla possesses a total length of 0.1 m of actin filaments²⁴; does the bundle's developmental programme rely on a constant amount of starting material? Why are hair bundles so vulnerable? Although lower vertebrates add hair cells throughout life, and the cells can be mitotically replaced in such animals and in birds^{25,26}, no hair cell has yet been shown to be capable of regenerating its hair bundle. Perhaps the precisely choreographed steps of bundle morphogenesis cannot be repeated in mature hair cells, so that damage due to ageing, acoustical trauma and ototoxicity leads, inevitably, to a decline in our

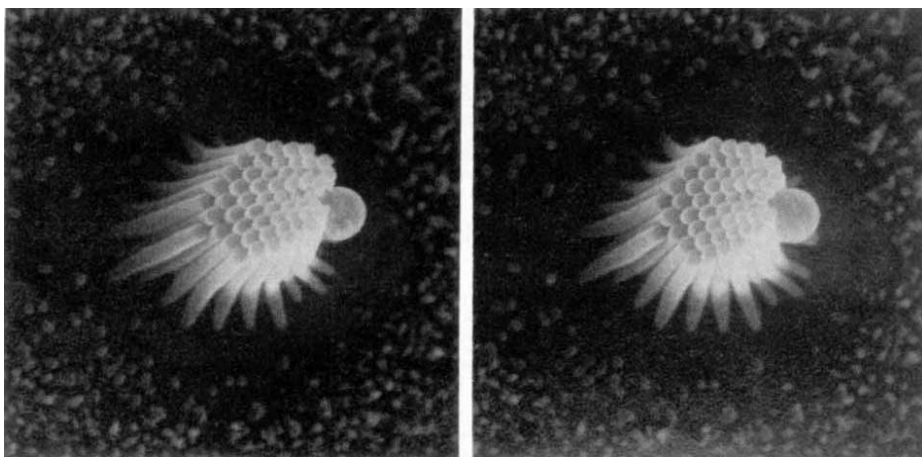


FIG. 2 A hair bundle of the bullfrog sacculus extends from the circular apical surface of a hair cell and comprises about 50 stereocilia and a single kinocilium with a bulbous termination. Magnification, $\times 9,000$. Gradation in stereociliary lengths, from $3 \mu\text{m}$ at the bundle's left edge to $7 \mu\text{m}$ at the right, confers a plane of mirror symmetry on the structure. Deflexion of the bundle's tip in the positive direction (to the right) is excitatory, whereas stimulation in the negative direction (to the left) is inhibitory. Fixation and critical-point drying of the specimen reduced the diameters of the stereocilia and caused bowing of their basal insertions, which is not observed in living cells^{11,13}. This stereo-pair scanning electron micrograph is meant to be viewed with divergent visual axes.

hearing capacity. How, then, are structural elements within the bundle renewed? Although it is implausible that individual actin monomers endure for a human lifetime, it is almost as difficult to see how they are replaced. Do new monomers substitute for exhausted ones by nick substitution, with re-annealing of the filament on both sides? Do new actin filaments grow alongside the existing ones, which somehow whither away? Or are whole stereocilia replaced, for example by addition of immature ones at the short edge of the bundle and progressive migration across the cellular surface to a subduction zone near the kinocilium?

Mechano-electrical transduction by hair cells

Mechanical stimuli applied to a hair bundle elicit electrical responses by regulating, or gating, the opening and closing of mechanically sensitive transduction channels³⁴. When a bundle is deflected by a probe attached at its tip (Fig. 3; ref. 27), the hair cell's response depends upon the direction and magnitude of the stimulus²⁸⁻³¹. In an unstimulated cell, 15% of the transduction channels are already open; the resting potential of such a cell, about -60 mV, is therefore determined in part by the inward flow of the transduction current. A positive stimulus, which displaces the bundle towards its tall edge, opens additional channels; the influx of cations depolarizes the cell by as much as tens of millivolts. A contrary stimulus, towards the short edge of the bundle, shuts those transduction channels open at rest and hyperpolarizes the cell. Stimuli at right angles to the bundle's axis of mirror symmetry, however, produce no change from the resting potential²⁸. Hair cells respond only to stimulus components parallel to the axis of morphological symmetry; an oblique stimulus accordingly elicits a response proportional to its vectorial projection along that axis.

The relation between a bundle's deflexion and the ensuing electrical response is in every instance sigmoidal (Fig. 4, refs 27-32). Although displacement-response relations vary in detail from cell to cell and from species to species, a consistent and striking feature is their steepness: about 90% of the response range corresponds to a deflexion of the bundle by only 50-120 nm at its tip, or through an angle of about $\pm 1^\circ$. In ordinary use, therefore, a hair bundle probably moves by less than the diameter of a single constituent stereocilium.

The threshold of hearing occurs at a sinusoidal hair-bundle deflexion near ± 0.3 nm (refs 12 and 33); this motion of $\pm 0.003^\circ$ corresponds to a displacement of the pinnacle of the Eiffel Tower by only a thumb's breadth. Threshold stimulation evokes a receptor potential of $\sim 100 \mu\text{V}$ (ref. 12), which evidently suffices for reliable synaptic transmission. From the stiffness of the hair bundle it is possible to calculate that free-standing bundles should exhibit brownian motion with a root-mean-square amplitude of 2 nm; interferometric observation confirms this inference³⁴. Although the auditory system is capable of improving its signal-to-noise ratio to some extent by averaging responses over several cycles, the sensitivity of our hearing is ultimately limited because fainter sounds are drowned out by the thermal clatter of the components of the ear^{35,36}.

Although they have not been identified biochemically, the mechano-electrical transduction channels of hair cells have been characterized by biophysical experiments. Pulses of transduction current observed in hair cells of the chick³⁷ may represent the opening of individual channels; if so, the single-channel conductance is 50 pS. Variance analysis indicates a conductance less than half as large in the hair cells of the frog³⁸. Transduction channels are relatively non-selective cation-passing pores^{39,40}, but K^+ , the ion present at the highest concentration in the endolymph around the hair bundle, normally carries most of the transduction current. Because small organic cations can support measurable current, the pore of the transduction channel must be at least 0.7 nm in diameter³⁹. Divalent cations exhibit a high permeability in this channel⁴⁰; the fact that Ca^{2+} traverses the channel less efficiently than K^+ (ref. 39) suggests that divalent ions have such a high affinity for the pore that they linger there

several times longer³. Aminoglycoside antibiotics, whose most deleterious side effect is ototoxicity exerted against hair cells, also enter and obstruct the channels^{40,41}.

Transduction channels are few in number. The steps in transduction current observed in chick hair cells, assuming they result from the opening of individual channels, suggest that these cells have under 20 channels apiece^{37,40}. Analysis of the variance in transduction current³⁸, as well as measurement of the forces involved in channel gating³², indicates that a hair cell from the bullfrog's sacculus has 50-100 channels. Because there are a comparable number of stereocilia in the hair bundle of such a cell, and as the magnitude of the response is roughly proportional to the number of stereocilia remaining in a microdissected bundle⁸, there may be only one active transduction channel per stereocilium. Together with the absence of high-affinity ligands with which to label the channels, this scarcity of channels accounts for the dearth of biochemical information about them.

Mechano-electrical transduction channels

What, then, opens mechano-electrical transduction channels? Although displacement *per se* may be measured with devices ranging from yardsticks to interferometers, the usual practice in engineering is to infer the magnitude of a displacement indirectly by detecting the strain that a force produces in some elastic material. A hair cell's sensitivity is ordinarily measured in terms of hair-bundle displacement, but there are several indications that mechano-electrical transduction in fact involves a biological strain gauge. The bundle seems to contain elements, the gating springs, the tension of which controls the opening and closing of transduction channels.

First, in measurements of stiffness of a hair bundle, it is possible to detect a component that is associated with mechano-electrical transduction. A bundle is stiffer along its axis of morphological symmetry than at a right angle; the extra stiffness, moreover, declines as a cell adapts to a protracted mechanical stimulus¹⁸. The evidence suggests that a portion of the work done in deflecting a bundle goes into the gating springs. These springs collectively contribute $\sim 50\%$ of the stiffness that opposes the bundle's flexion^{18,32}; as a result, the transducer efficiently harvests the energy supplied when the bundle is stimulated.

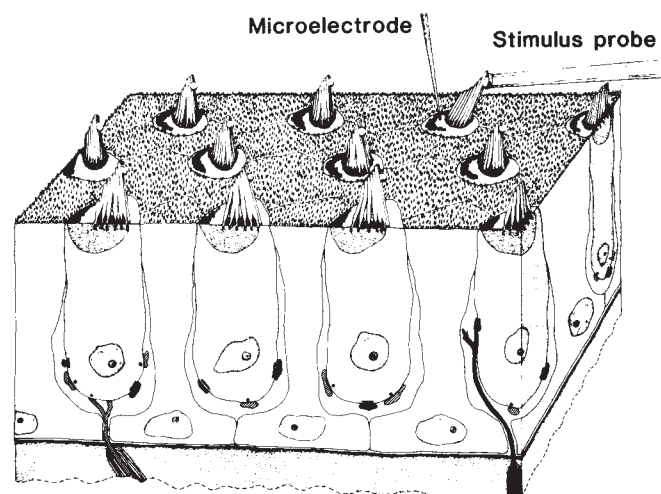


FIG. 3 Experimental procedure for recording electrical responses from a hair cell *in vitro*²⁷. Hair cells from the bullfrog sacculus lie in an epithelial sheet, separated by supporting cells with apices covered by microvilli. Afferent and efferent synapses stud the hair cells' basolateral surfaces. A microelectrode is used to measure the cell's receptor potential or, under voltage-clamp conditions, its transduction current. A glass probe applied to the hair bundle's distal end enables the experimenter to provide a mechanical stimulus of any desired orientation, amplitude and waveform. For measurements of bundle stiffness, the stimulus probe is replaced by a flexible glass fibre positioned at a right angle to the probe illustrated here.

The time course of channel gating provides a second line of evidence for gating springs. The response latency is so brief—in the order of microseconds—that gating is probably direct, rather than operating through a second messenger⁴². Moreover, larger step stimuli elicit electrical responses that (up to a saturating level) are not only larger, but also rise faster^{31,43}. This behaviour favours a kinetic scheme in which mechanical force controls the rate constants for channel gating. If the mechanical energy from a stimulus were stored in a spring attached to the channel's gate, the rates of channel opening and closing would be determined by the probability that the energy content of the spring exceeds the transition-state energy for channel opening or closing. Because stimuli primarily affect the opening rate constant^{31,43}, the rate-limiting, transition-state configuration of the transduction channel probably resembles the open state³.

The third line of evidence for gating springs can be introduced through a homely analogy. Imagine tying an elastic cord to the knob of a closed door, then measuring the cord's stiffness by applying a known force at its free end and measuring its consequent extension. A similar estimate of the stiffness would result if the door were instead held wide open. Suppose, however, that the experiment were repeated with the door initially closed, but free to open. As one exerted the usual force on the cord, its free end would move further as the door swung open, in the direction of the force. The experimenter might conclude that the cord is less stiff over the range of displacements in which the door opens.

The hair bundle exhibits such an effect: the bundle is ~40% more flexible near its resting position than at the extremes of positive or negative displacement (Fig. 4; refs 32 and 44). Three lines of evidence associate this augmented compliance with the gating of transduction channels: the range of bundle displacements over which the compliance increases coincides with that over which channels open and close; the compliance is blocked by aminoglycoside drugs that obstruct transduction channels; and the bundle's position of enhanced compliance comigrates during adaptation with the position of mechanosensitivity. Quantitative analysis of the gating compliance confirms that there are only one to a few transduction channels per stereocilium³². The results also suggest that individual gating springs are about as stiff as single protein filaments. The gate of the transduction channel evidently swings through 4 nm upon opening, which is opposed by a restoring force of about 2 pN.

The existence of gating compliance implies a reciprocal interaction between the forces on the hair bundle and the gating of transduction channels. Not only does force applied to the bundle influence the probability of a channel being open, but the channel's status affects the force exerted by the bundle. This observation therefore buttresses the argument that transduction channels are mechanically gated by elastic elements: if a second messenger intervened in the transduction process between the force-sensitive apparatus and the ion channel, gating compliance would not be expected. Along with the responses of motor proteins, gating compliance constitutes one of the few mechanically measured signals of a molecular conformational change of the sort invoked to explain the activities of enzymes and ion channels.

What, and where, are the gating springs?

Despite the evidence that stimulus forces are directly communicated to transduction channels by gating springs, the identity of these elements remains uncertain. Gating springs might, for example, be cytoskeletal components of the hair bundle or integral parts of the transduction-channel molecules themselves. The plasma membrane is also an appropriately elastic structure that might constitute a gating spring, in which case transduction channels could prove to be stretch-activated⁴⁵. Lacking a histochemical or other morphological marker for the transduction channels, we must infer the location of the gating springs from indirect evidence. Unfortunately, the two relevant physiological

experiments to date provide seemingly contradictory results.

From the polarities of extracellularly and intracellularly recorded responses, we know that the cationic current that enters a hair cell through transduction channels must cross the apical plasma membrane⁴⁶, which includes the membrane covering the stereocilia and kinocilium, as well as that on the flattened cellular surface. To reach the apical membrane, the current flows through the saline solution around the bundle; in so doing, it must inevitably produce a tiny potential drop across the resistance that this medium affords. Measurement of the extracellular potentials about a stimulated hair bundle points to a current sink at the top of the bundle, near the tips of the stereocilia (Fig. 5; ref. 47). The signals are of the magnitude anticipated from the size of the transduction current and the resistivity of the medium. The responses also display the expected sensitivities to membrane potential, to an ototoxic antibiotic, and to the direction and amplitude of stimulation.

The current-sink experiment provides evidence that transduction channels lie near the tips of stereocilia. This interpretation, however, rests on the presumption that the extracellular flow of current toward the transduction channels is not severely distorted by the presence of the hair bundle. This assumption seems reasonable because electron microscopy indicates that about half of the volume within a bundle is extracellular space to which large proteinaceous markers have access (R. A. Jacobs and A. J. H., unpublished observation) and through which ionic current should readily flow.

Transduction current entering channels near the stereociliary tips would have to flow axially down the stereocilia before it could effect a change in the membrane potential of the soma, and hence in the release of synaptic transmitter. Unless the internal resistivity of the stereociliary interior is unusually high, however, the decrement in steady-state responses should be only a few per cent⁴⁷. Because of current flow across the stereociliary membrane capacitance, responses to high-frequency stimulation could be more severely attenuated. Perhaps in part to counter this problem, however, the hair bundles of organs responsive to high-frequency sounds are systematically shorter than those in receptors for low-frequency stimuli^{19,48,49}. To cite extreme cases, the bat's cochlear receptors for 83 kHz ultrasound have 0.8 μm stereocilia¹⁵, whereas receptors for low-frequency angular acceleration are endowed with stereocilia over 90 μm in length⁵⁰.

An alternative method for locating transduction channels relies on the fact that these channels can pass divalent cations. On the assumption that intracellular accumulation of Ca^{2+} following mechanical stimulation reveals the site of transduction, a Ca^{2+} -sensitive fluorescent dye, fura-2, has been used to follow the appearance of Ca^{2+} in the cytoplasm of hair cells⁵¹. The largest fluorescence signals are observed in the apical cytoplasm,

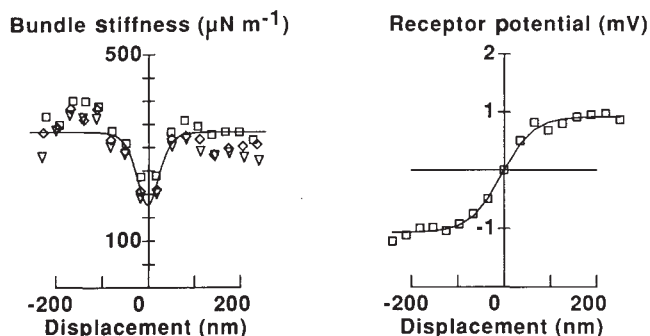


FIG. 4 Gating compliance is a reduction in the hair-bundle stiffness (left) measured by applying known forces to a bundle with a flexible glass fibre, then calculating from the fibre's flexion the stiffness of the attached bundle⁴⁴. The fact that compliance is greatest over the range of hair-bundle positions where transduction is most sensitive (right) suggests that the compliance results from the opening and closing of transduction channels.

immediately beneath the hair bundle, suggesting that transduction channels lie at the bases of the stereocilia or in the adjacent membrane of the flattened cellular apex.

Although the Ca^{2+} -detection experiment provides a useful second method for seeking the site of transduction, the indirect nature of the measurement introduces three complications into interpretation of the data. First, because the average Ca^{2+} ion would be expected to diffuse several micrometres while bound to the dye^{52,53}, the results do not pinpoint the site of Ca^{2+} entry. Second, the experiment shows where Ca^{2+} accumulates after continuous stimulation for several hundred milliseconds; but as the membrane-permeable form of fura-2 is used⁵¹, the result could reflect the gradual accumulation of Ca^{2+} in vesicular organelles around the cuticular plate⁴⁹. Third, the thin stereocilia of a hair bundle, even when well filled with dye, provide a much weaker fluorescence signal than does the nearby soma. The published data therefore do not exclude the possibility that Ca^{2+} enters stereocilia near their tips, then diffuses to a site in the soma where it produces a more prominent optical signal than within the bundle.

Ultrastructural analysis of hair cells has failed to resolve the lingering uncertainty about the site of transduction. Although freeze-fracture studies demonstrate intramembrane particles on the apical cellular surface, there is no evidence for specializations at either the bases or the tips of stereocilia^{54,55}. Scanning and transmission electron microscopy has, however, put forward a specific type of inter-stereociliary connection as a candidate to be the gating spring. This tip link is a single strand joining the distal end of almost every stereocilium to the side of the longest adjacent process (Fig. 6; refs 56–58). Not only are tip links well situated to sense hair-bundle motion, but their orientation could account for the vectorial sensitivity of transduction. Pushing a bundle in the positive direction should elongate the tip links and, if they are the gating springs, promote channel opening. Oppositely directed stimuli should slacken the links and close the channels. Stimulation at a right angle to the bundle's plane of symmetry would be expected to elicit little response, for the links invariably run in a direction parallel with that plane.

Identification of the tip links with the gating springs is alluring: the links occur in every hair bundle for which data are available, and their position at the bundle's top would accord with measurement of a transduction-current sink there. There is, nevertheless, no direct evidence that the tip links are the gating springs. It would be very desirable, for example, to confirm that the links are elongated during excitatory stimulation, shortened upon inhibitory stimulation, and unaffected by stimuli that elicit no change in the transduction current. Until such demonstrations have been accomplished, or until the tip links have been shown to be unnecessary (perhaps by being enzymatically digested while transduction survives), the hypothesis will remain only that.

Adaptation

Despite the precision with which it forms, a hair bundle is unlikely to develop so accurately that the sensitive transduction apparatus is perfectly poised at its position of greatest mechanosensitivity. Some mechanism would seem necessary to compensate for developmental irregularities, as well as for environmental changes, by resetting the gating springs to accord with the bundle's resting position⁵⁹. An adaptation process that continuously adjusts the bundle's range of operation does just that, maintaining high sensitivity to transient stimuli while in certain hair cells rejecting static inputs a million times as large⁶⁰.

Adaptation manifests itself as a progressive decrease in the transduction current following the onset of a protracted hair-bundle deflection⁶¹. The application of test stimuli during the course of adaptation reveals that the process differs from desensitization: transduction persists with normal sensitivity, but the position at which the transducer is most sensitive

migrates from the bundle's resting position towards that at which the bundle is held. Adaptation occurs on a timescale three orders of magnitude slower than mechanoelectrical transduction: the time constant of adaptation to a half-saturating stimulus is ~ 30 ms with endolymph bathing the hair bundle⁶¹. Both the rate and the extent of adaptation increase with increasing concentration of Ca^{2+} in the fluid contacting the apical cellular surface^{61,62}.

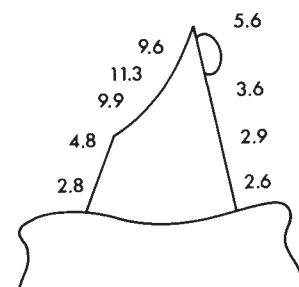
How does adaptation occur? Because the hair bundle becomes measurably less stiff as adaptation proceeds¹⁸, the process evidently involves a reduction of the tension in the gating springs. This could come about in several ways; the cuticular plate could warp so as to adjust the interrelation of the stereocilia, for example, or the individual stereociliary cytoskeletons could change dimensions. If the tip links are the gating springs, the most likely possibility is that the anchoring points for these links are repositioned in a negative-feedback cycle (Fig. 7; ref. 63). Suppose that a gating spring is kept tense by an attached motor protein, for example a myosin I molecule similar to the protein of relative molecular mass 110,000 found in microvilli⁶⁴, crawling up a cytoskeletal actin filament. When a stimulus step increases the tension in the spring, the associated transduction channel spends a larger fraction of its time open. The increased cation influx includes Ca^{2+} , which accumulates in the stereocilium and interacts with the motor protein, decreasing its upward force. The gating spring then shortens by pulling down the motor molecule; when the spring reaches its resting tension, closure of the channel reduces the Ca^{2+} influx to its original level, restoring a balance between the upward force produced by the protein and the downward tension in the spring. In this model, Ca^{2+} would function as a second messenger in hair-cell adaptation rather as it does in the light-adaptation process of photoreceptors⁶⁵.

According to this model for adaptation⁶³, the diminution in stiffness of the hair bundle during adaptation would be expected as a consequence of declining tension in the gating springs¹⁸. Another result consistent with the model is that extensive depolarization of a cell, which would be expected to diminish the influx of Ca^{2+} through transduction channels, elicits motion of the hair bundle in the negative direction, towards its short edge⁶². This response is to be expected if the reduced influx of Ca^{2+} causes the motor protein to increase the tension in the gating springs.

Cellular contributions to frequency selectivity

Although they must contend with stimuli of very low energy content, hair cells enjoy the advantage that these inputs are often periodic in nature. If the significant part of stimulus consists of a regular signal, such as the sinusoidal pressure due to a pure-tone acoustical stimulus, a detection system can increase its signal-to-noise ratio by selectively amplifying the response to a relevant frequency. At least two mechanisms are now known that accomplish this task on a cellular level.

FIG. 5 Localization of the site of mechanoelectrical transduction by determination of extracellular current flow. While a hair cell is mechanically stimulated, extracellular potentials are recorded at various sites about the hair bundle⁴⁷. Flowing through saline solution with a resistivity of $1 \Omega \text{ m}$, a typical transduction current of 200 pA would be expected to produce signals of the observed magnitudes (shown as average peak-to-peak values in microvolts). The strong response at the bundle summit indicates that mechanoelectrical transduction occurs there, near the stereociliary tips.



The first cellular frequency filter occurs prior to mechanoelectrical transduction: the mechanical properties of each hair bundle conspire to tune it to a particular frequency. A tuning fork's resonant frequency depends upon the flexibility and mass of its tines; in a similar way, the natural frequency of a hair bundle depends upon its mechanical properties. The elastic elements that restore a bundle to its upright, resting position include—and are probably dominated by—the actin cables at the bases of the stereocilia^{12,13}. Because the bundle moves through an aqueous medium of comparable density, the relevant mass in the bundle's tuning includes that of a volume of water viscously dragged along by the moving bundle⁶⁶. Viscosity also heavily damps the motion: although a bundle responds optimally at a particular frequency, it differs from a tuning fork in being incapable of sustained oscillation after stimulation has ceased.

As noted above, the hair bundles of many auditory organs display a gradation in length along the frequency axis. In line with the tuning-fork mechanism of frequency selectivity, the bundles are relatively long on hair cells that respond to low-frequency acoustical, vibrational and accelerational stimuli⁵⁰, whereas the shortest bundles occur on receptors for the highest-frequency acoustic signals¹⁵. The stiffnesses of hair bundles are also inversely related to their lengths¹⁷.

The motion of hair bundles is difficult to simulate for several reasons: bundles are of a complex shape, have many internal degrees of mechanical freedom, occur at the surface between a fluid and a solid, and lie close enough together that their movements are influenced by those of their neighbours. Despite this hydrodynamic complexity, computer modelling^{66,67} shows that the tuning-fork mechanism can account for the observed frequency selectivity of cells with free-standing hair bundles^{22,23}, which are not attached to a tectorial membrane. Stereociliary length may also affect the tuning of cells whose hair bundles are inserted into a tectorial membrane⁶⁸, for these cells too exhibit an inverse relation between bundle length and characteristic frequency^{48,49}.

The second mechanism that tunes individual hair cells to specific frequencies is electrical in nature⁶⁹, and, in contrast to mechanical tuning, occurs subsequent to mechanoelectrical transduction. The membrane properties of certain hair cells in amphibians⁷⁰, reptiles⁷¹ and birds⁷² are such that the membrane potential tends to resonate at a particular frequency. This characteristic is demonstrated by injecting a pulse of current into a hair cell, whereupon the membrane potential undergoes an exponentially damped, sinusoidal oscillation⁷¹. When such a cell is stimulated with mechanical stimuli of constant amplitude, its transduction apparatus responds over a broad range of frequencies. Stimulation at the particular frequency at which a cell's membrane potential resonates when current is injected, however, evokes the largest receptor potential. Because the release of synaptic transmitter at the basolateral surface of the cell is controlled by membrane potential, the postsynaptic nerve fibres innervating the cell are most responsive to stimuli of the specific frequency⁷³.

The origin of electrical resonance has been determined by recordings from enzymatically isolated hair cells with the whole-cell and patch variants of the tight-seal electrode technique^{74,75}. Current carried into a hair cell through voltage-sensitive Ca^{2+} channels drives the depolarizing phase of an oscillation; the hyperpolarizing component results from current through Ca^{2+} -sensitive K^+ channels⁷⁰. Several factors establish such features of the resonance as its frequency and damping: the membrane capacitance, the numbers and kinetic properties of the Ca^{2+} and K^+ channels and the time course of Ca^{2+} removal from the membrane by diffusion, sequestration and extrusion^{74,76,77}.

It remains to be determined how hair cells become tuned to their appropriate frequencies^{69,77}. In the basilar papillae of reptiles and birds, for example, there is a continuous representation of best frequencies from one end of the organ to the other.

Some of the systematic variation in oscillation frequency accrues from differences in the gating kinetics of K^+ channels⁷⁴. Do these channels occur in several isoforms that are appropriately doled out to cells, or are channels of one or a few sorts post-translationally modified to adjust their kinetic properties? The behaviour of K^+ channels might instead be dominated by factors that control the rate of disappearance of Ca^{2+} from beneath the membrane, in which case the regulation of Ca^{2+} homeostasis will prove of primary importance in the establishment of tuning.

Cochlear frequency selectivity and force production

In the mammalian cochlea, pure-tone stimulation evokes travelling waves that propagate from the base of the basilar membrane towards its apex, peaking in amplitude at a specific point along the way⁷⁸. Each position along the basilar membrane is most sensitive to a particular frequency of stimulation; the characteristic frequencies are arranged in a monotonic order along the membrane, with high frequencies mapped to the cochlea's base, low frequencies to its apex. The sharpness of frequency discrimination is similar, whether measured in terms of basilar-membrane motion^{33,79}, hair-cell response⁸⁰, or eighth-nerve activity⁸¹. Complex acoustic stimuli, which can be resolved into multiple sinusoidal components, evoke a basilar-membrane disturbance that is (at least approximately) the superposition of the membrane's response to the constituent tones: the basilar membrane acts as a frequency analyser that apportions stimulus energy to the hair cells arrayed along its length according to a sound's content of various pure tones.

The tuning observed in the cochlea appears too sharp to result solely from the passive hydrodynamic and mechanical properties of the basilar membrane. Frequency selectivity might be enhanced through mechanical resonance by the tectorial membrane⁸², the gelatinous mass into which the hair-bundle tips of outer hair cells are inserted. Another possibility is that tuning is sharpened by force-producing elements that counter the damping effects of the cochlear fluids^{83–85}. The strongest evidence for such an active process is that cochleae can emit sounds, both spontaneously⁸⁶ and in response to acoustic stimulation⁸⁷ at frequencies as high as 62 kHz (ref. 88). Because the sound that spontaneously emerges from ears has an energy and statistical properties that cannot be explained by narrow-band

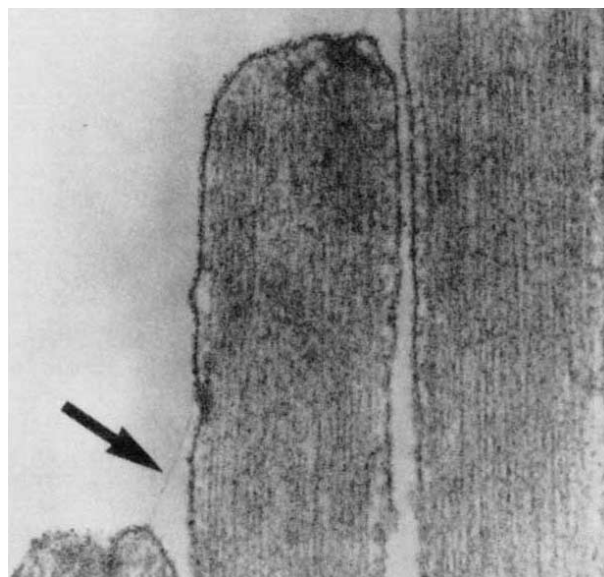


FIG. 6 A tip link, shown here in an electron micrograph of a hair cell from the bullfrog sacculus, is a thin fibre (arrow) that connects the end of one stereocilium to the side of the next. Magnification, $\times 115,000$. Such a link may be the gating spring that opens a transduction channel at one end or the other (or both). Evenly spaced actin filaments form the cytoskeletons of the stereocilia.

filtering of noise from within an animal⁸⁹, one must conclude that something in the ear actively moves, liberating energy in the form of mechanical vibrations.

What supplies this energy? Several lines of evidence implicate the outer hair cells of the mammalian cochlea as the elements that enhance frequency selectivity, and hence as the energy sources. These cells, which spiral in three rows alongside the single row of inner hair cells, make only token projections to the central nervous system through their afferent fibres. They receive a copious efferent innervation⁹⁰, however, whose activation decreases both cochlear sensitivity and frequency discrimination⁹¹. Pharmacological ablation of outer hair cells affects responsiveness even more profoundly⁹².

An isolated outer hair cell can display motility *in vitro*^{93,94}: when stimulated electrically, it shortens when depolarized, elongates when hyperpolarized^{95,96}. The energy for the effect could be derived from the experimentally imposed electrical field, rather than from hydrolysis of an energy-rich substrate^{4,96,97}. Consistent with this, such a cell displays a component of capacitive current that reflects the motion of voltage-sensitive molecules in the cell membrane⁹⁷; these moving molecules might distort the elastic cytoskeleton⁹⁸.

Despite the demonstrations that outer hair cells can shorten, several reasons remain for scepticism that such contraction constitutes the motor for cochlear tuning and emission. First, because outer hair cells are orientated nearly perpendicular to the basilar membrane in some cochleae, including those of humans¹, changes in the cells' length would be expected to apply forces principally along hair bundles, rather than at a right angle, along the bundles' axes of mechanosensitivity. For that reason, the motion relevant to tuning might prove to be tipping of the hair cells' apical ends⁹⁹. Second, motility based upon voltage-sensitive membrane components should suffer attenuation at high frequencies because of the membrane's capacitance. The outer hair cell's membrane time constant of 0.13–3 ms (refs 100, 101) corresponds to a corner frequency between 50 Hz and 1.3 kHz. Although voltage-driven mechanical responses could occur at higher frequencies, they would become progressively less efficient. Third, because hydrodynamic drag is proportional to the velocity of a moving object, the energetic cost of cell-body motion would rise with increasing frequency. Fourth, the electromechanical gain of isolated outer hair cells, $2 \mu\text{m V}^{-1}$ (ref. 101) at the resting potential of -70 mV (ref. 100), seems too small for motility to exert a significant effect on transduction in response to receptor potentials near threshold¹⁰¹. Finally, acoustic emissions are not confined to the mammalian cochlea, the sole domain of morphologically specialized outer hair cells; evoked emissions occur in birds¹⁰² and spontaneous emission even in amphibians^{103,104}. The energy-producing process, like

other features of mechanoelectrical transduction, seems to be common to hair cells throughout the vertebrates. It remains possible, for example, that all hair cells can display the shortening response under appropriate conditions.

Another site of force production that might contribute to frequency tuning is the hair bundle, which may prove to be more than a recipient of stimuli. In addition to displaying passive elasticity, some bundles are capable of actively exerting forces against other objects. The bundles of electrically resonant hair cells display back-and-forth motion in phase with the electrical signals; a bundle moves in the positive direction as the membrane depolarizes^{12,18}. In addition, following displacement in the positive direction by a flexible fibre, a bundle can transiently push back against the probe, producing a rebound motion^{18,32}. There are no data on the mechanism of force production by hair bundles, nor is there evidence that bundles can generate forces at the very high frequencies at which sharp frequency selectivity and acoustic emission are observed.

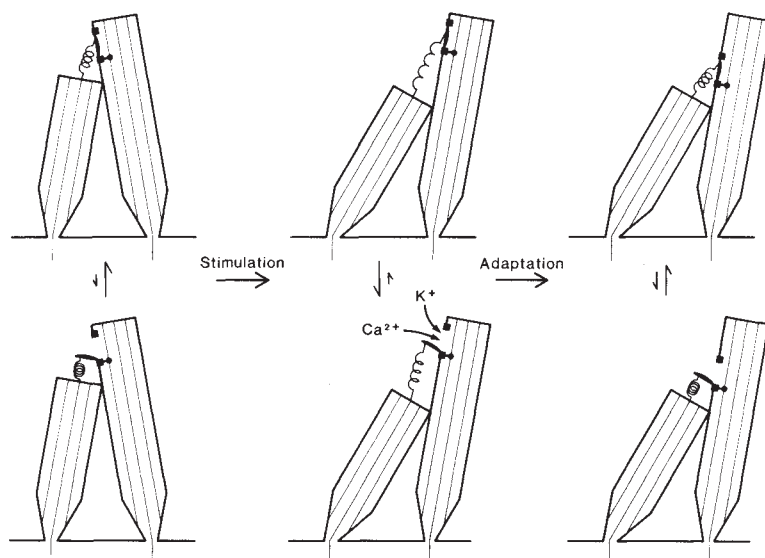
Why is transduction direct?

In contrast to hair cells, many other sensory-receptor cells—vertebrate photoreceptors¹⁰⁵ and olfactory neurons¹⁰⁶, for example—employ cyclic nucleotides or other second messengers in transduction. Two advantages stem from this strategy: the enzymatic apparatus that generates a second messenger provides amplification, and feedback within the metabolic pathway permits gain control such as adaptation or desensitization.

What is the selective advantage of transduction without the intervention of a second messenger? The answer probably lies in the speed of response; hair cells operate much more quickly than do other sensory-receptor cells of the vertebrate nervous system, and indeed, faster than neurons themselves. Transduction by hair cells must be rapid to deal with the frequencies of biologically relevant stimuli. The behaviour of sound in air and the physical dimensions of animals, particularly those of sound-emitting and -absorbing organs such as vocal cords and eardrums, make frequencies in the kilohertz range optimal for auditory communication. Much lower frequencies are inefficiently captured by animals of moderate size; much higher frequencies propagate poorly through air.

The localization of sound sources, one of the most important roles of hearing, sets even more stringent limits on the speed of transduction. If a sound source lies directly to one side of an animal, an emitted sound will reach the nearer ear somewhat sooner than the farther; for a human, this maximal delay is about $700 \mu\text{s}$ (ref. 107). Both humans¹⁰⁸ and owls¹⁰⁹ can locate sound sources on the basis of much smaller temporal delays, less than $20 \mu\text{s}$. For this to occur, hair cells must be capable of detecting acoustical waveforms with microsecond-level resolu-

FIG. 7 A model for transduction and adaptation by hair cells⁶³. Each stereocilium has near its tip a transduction channel attached to a gating spring. When the bundle is in its resting position (left), the channel rattles between its closed (upper) and open (lower) configurations about one thousand times per second; as suggested by the relative sizes of the reaction-rate arrows, however, the channel is shut most of the time. Immediately after a positive stimulus is applied (centre), shear between the stereocilia increases tension in the gating spring and promotes channel opening; entry of cations such as K^+ and Ca^{2+} then depolarizes the cell. As adaptation occurs during a few tens of milliseconds (right), repositioning of the gating spring's end, possibly by a motor protein descending the actin cytoskeleton of the stereocilium, reduces tension in the gating spring and allows channel closure.



ution; even more remarkably, this temporal information must be transmitted along axons whose individual signalling rate is less than 1 kHz. The unusual innervation ratio of the mammalian cochlea, with more than 20 axons innervating a single inner hair cell in some instances⁹⁰, may be required for data transmission at a suitable rate.

Active echo-location, the hunting strategy employed by some bats and whales, places extreme demands on the speed of auditory transduction. To detect potential prey, let alone to resolve surface features that are useful in its identification, a bat must emit sound with a wavelength comparable to the dimensions of the target. A 3 mm insect, for example, efficiently reflects sound only at frequencies near and above 100 kHz. Chiropterans and cetaceans, whose auditory responses extend to at least 150 kHz, exploit the direct nature of hair cell transduction to the fullest.

The responsiveness of hair cells to high frequencies of stimulation implies that transduction channels are very rapidly gated. Even in frogs sensitive to relatively low frequencies, a stimulus of moderate intensity causes a response with a time constant of only 80 μ s at 24°C. The activation enthalpy of the channel-opening reaction, about 50 kJ mol⁻¹ (ref. 43), implies that such channels would open with a time constant as small as 30 μ s at

37°C. To respond to frequencies in excess of 100 kHz, mammalian hair cells might have gating rates as much as an order of magnitude faster.

The rate of gating of the mechanoelectrical transduction channel is remarkable by comparison with the pace of spontaneous rearrangement of protein molecules. A structural domain of the size expected for the transduction-channel gate probably diffuses on a timescale of approximately a millisecond¹¹⁰. There is, however, no inherent difficulty with a much faster-moving gate: breathing motions of a protein are driven by thermal energy, but channels can be gated by externally supplied energy. A sodium channel, for example, is thought to be opened when an electrical field causes sliding motions of four charged, α -helical segments of the channel protein¹¹¹. This relaxation, which involves the relocation of about 90 amino-acid residues, can occur with a time constant of 10 μ s (ref. 112). If a hair cell's transduction channel opens when a force directly displaces a gate from the ion-passing pore, the mechanical energy supplied by the stimulus can account for the channel's extraordinary—and most useful—speed of operation. □

A. J. Hudspeth is at the Department of Cell Biology and Neuroscience, University of Texas Southwestern Medical Centre, 5323 Harry Hines Boulevard, Dallas, Texas 75235, USA.

1. Retzius, G. *Das Gehörorgan der Wirbelthiere*, Volume II (Samson and Wallin, Stockholm, 1884).
2. Watanuki, K. & Schuknecht, H. F. *Arch. Otolaryngol.* **102**, 583–588 (1976).
3. Howard, J., Roberts, W. M. & Hudspeth, A. J. *Rev. Biophys. biophys. Chem.* **17**, 99–124 (1988).
4. Roberts, W. M., Howard, J. & Hudspeth, A. J. *Rev. Cell Biol.* **4**, 63–92 (1988).
5. Flock, Å. & Cheung, H. C. *J. Cell Biol.* **75**, 339–343 (1977).
6. De Rosier, D. J., Tilney, L. G. & Egelman, E. *Nature* **287**, 291–296 (1980).
7. Tilney, L. G., DeRosier, D. J. & Mulroy, M. J. *J. Cell Biol.* **86**, 244–259 (1980).
8. Hudspeth, A. J. & Jacobs, R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1506–1509 (1979).
9. Flock, Å., Bretschger, A. & Weber, K. *Hearing Res.* **6**, 75–89 (1982).
10. Shepherd, G. M. G., Barres, B. A. & Corey, D. P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4973–4977 (1989).
11. Tilney, L. G., Saunders, J. C., Egelman, E. & DeRosier, D. J. *Hearing Res.* **7**, 181–197 (1982).
12. Crawford, A. C. & Fettiplace, R. *J. Physiol., Lond.* **364**, 359–379 (1985).
13. Howard, J. & Ashmore, J. F. *Hearing Res.* **23**, 93–104 (1986).
14. Miller, M. R. *J. Morphol.* **156**, 381–418 (1978).
15. Bruns, V. & Goldbach, M. *Anat. Embryol.* **161**, 51–63 (1980).
16. Engström, H., Ades, H. W. & Hawkins, J. E., Jr. *J. Acoust. Soc. Am.* **34**, 1356–1363 (1962).
17. Strelhoff, D. & Flock, Å. *Hearing Res.* **15**, 19–28 (1984).
18. Howard, J. & Hudspeth, A. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3064–3068 (1987).
19. Tilney, L. G. & Saunders, J. C. *J. Cell Biol.* **96**, 807–821 (1983).
20. Tilney, L. G., Tilney, M. S., Saunders, J. S. & DeRosier, D. J. *Dev. Biol.* **116**, 100–118 (1986).
21. Tilney, L. G. & DeRosier, D. J. *Dev. Biol.* **116**, 119–129 (1986).
22. Frishkopf, L. S. & DeRosier, D. J. *Hearing Res.* **12**, 393–404 (1983).
23. Holton, T. & Hudspeth, A. J. *Science* **222**, 508–510 (1983).
24. Tilney, L. G. & Tilney, M. S. *J. Cell Biol.* **107**, 2563–2574 (1988).
25. Corwin, J. T. & Cotanche, D. A. *Science* **240**, 1772–1774 (1988).
26. Ryals, B. M. & Rubel, E. W. *Science* **240**, 1774–1776 (1988).
27. Hudspeth, A. J. & Corey, D. P. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2407–2411 (1977).
28. Shotwell, S. L., Jacobs, R. & Hudspeth, A. J. *Ann. N. Y. Acad. Sci.* **374**, 1–10 (1981).
29. Russell, I. J., Richardson, G. P. & Cody, A. R. *Nature* **321**, 517–519 (1986).
30. Ohmori, H. *J. Physiol., Lond.* **387**, 589–609 (1987).
31. Fettiplace, R. & Crawford, A. C. in *Cochlear Mechanisms: Structure, Function and Models* (eds Wilson, J. P. & Kemp, D. T.) 99–106 (Plenum, New York, 1989).
32. Howard, J. & Hudspeth, A. J. *Neuron* **1**, 189–199 (1988).
33. Sellick, P. M., Patuzzi, R. & Johnstone, B. M. *J. Acoust. Soc. Am.* **72**, 131–141 (1982).
34. Denk, W., Webb, W. W. & Hudspeth, A. J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5371–5375 (1989).
35. de Vries, H. *Physica* **14**, 48–60 (1948).
36. Harris, G. G. *J. Acoust. Soc. Am.* **44**, 176–186 (1968).
37. Ohmori, H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1888–1891 (1984).
38. Holton, T. & Hudspeth, A. J. *J. Physiol., Lond.* **375**, 195–227 (1986).
39. Corey, D. P. & Hudspeth, A. J. *Nature* **281**, 675–677 (1979).
40. Ohmori, H. *J. Physiol., Lond.* **359**, 189–217 (1985).
41. Kroese, A. B. A., Das, A. & Hudspeth, A. J. *Hearing Res.* **37**, 203–218 (1989).
42. Corey, D. P. & Hudspeth, A. J. *Biophys. J.* **26**, 499–506 (1979).
43. Corey, D. P. & Hudspeth, A. J. *J. Neurosci.* **3**, 962–976 (1983).
44. Hudspeth, A. J., Roberts, W. M. & Howard, J. in *Cochlear Mechanisms: Structure, Function and Models* (eds Wilson, J. P. & Kemp, D. T.) 117–124 (Plenum, New York, 1989).
45. Guharay, F. & Sachs, F. *J. Physiol., Lond.* **352**, 685–701 (1984).
46. Corey, D. P. & Hudspeth, A. J. *J. Neurosci.* **3**, 942–961 (1983).
47. Hudspeth, A. J. *J. Neurosci.* **2**, 1–10 (1982).
48. Wright, A. *Hearing Res.* **13**, 89–98 (1984).
49. Lim, D. J. *Hearing Res.* **22**, 117–146 (1986).
50. Helmholtz, H. L. F. *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, (Ellis, A. J., trans.) 138 (Dover, New York, 1954).
51. Ohmori, H. *J. Physiol., Lond.* **399**, 115–137 (1988).
52. Baylor, S. M. & Hollingworth, S. J. *J. Physiol., Lond.* **403**, 151–192 (1988).
53. Klein, M. G., Simon, B. J., Szucs, G. & Schneider, M. F. *Biophys. J.* **53**, 971–988 (1988).
54. Jahnke, K. *Acta Otolaryngol. Suppl.* **336**, 5–40 (1975).
55. Gulley, R. L. & Reese, T. S. *Anat. Rec.* **189**, 109–123 (1977).
56. Pickles, J. O., Comis, S. D. & Osborne, M. P. *Hearing Res.* **15**, 103–112 (1984).
57. Furness, D. N. & Hackney, C. M. *Hearing Res.* **18**, 177–188 (1985).
58. Osborne, M. P., Comis, S. D. & Pickles, J. O. *Hearing Res.* **35**, 99–108 (1988).
59. Hudspeth, A. J. *Trends Neurosci.* **6**, 366–369 (1983).
60. Narins, P. M. & Lewis, E. R. *J. Acoust. Soc. Am.* **76**, 1384–1387 (1984).
61. Eatock, R. A., Corey, D. P. & Hudspeth, A. J. *J. Neurosci.* **7**, 2821–2836 (1987).
62. Assad, J. A., Hachonen, N. & Corey, D. P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2918–2922 (1989).
63. Howard, J. & Hudspeth, A. J. *Discuss. Neurosci.* **4**, 138–145 (1987).
64. Collins, J. H. & Borysenko, C. W. *J. Biol. Chem.* **259**, 14128–14135 (1984).
65. Matthews, H. R., Murphy, R. L. W., Fain, G. L. & Lamb, T. D. *Nature* **334**, 67–69 (1988).
66. Freeman, D. M. & Weiss, T. F. *Hearing Res.* **35**, 201–207 (1988).
67. Weiss, T. F. & Leong, R. *Hearing Res.* **20**, 157–174 (1985).
68. Strelhoff, D., Flock, Å. & Minser, K. E. *Hearing Res.* **18**, 169–175 (1985).
69. Fettiplace, R. *Trends Neurosci.* **10**, 421–425 (1987).
70. Lewis, R. S. & Hudspeth, A. J. *Nature* **304**, 538–541 (1983).
71. Crawford, A. C. & Fettiplace, R. *J. Physiol., Lond.* **312**, 377–412 (1981).
72. Fuchs, P. A., Nagai, T. & Evans, M. G. *J. Neurosci.* **8**, 2460–2467 (1988).
73. Crawford, A. C. & Fettiplace, R. *J. Physiol., Lond.* **306**, 79–125 (1980).
74. Art, J. J. & Fettiplace, R. *J. Physiol., Lond.* **385**, 207–242 (1987).
75. Hudspeth, A. J. & Lewis, R. S. *J. Physiol., Lond.* **400**, 237–274 (1988).
76. Ashmore, J. F. & Attwell, D. *Proc. Roy. Soc. B226*, 325–344 (1985).
77. Hudspeth, A. J. & Lewis, R. S. *J. Physiol., Lond.* **400**, 275–297 (1988).
78. von Békésy, G. *Experiments in Hearing* (Wever, E. G., trans.) 485–510 (McGraw-Hill, New York, 1960).
79. Khanna, S. M. & Leonard, D. G. B. *Science* **215**, 305–306 (1982).
80. Russell, I. J. & Sellick, P. M. *J. Physiol., Lond.* **384**, 261–290 (1987).
81. Kiang, N. Y.-S. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT, 1965).
82. Zwillocki, J. J. & Kletsky, E. J. *Hearing Res.* **2**, 501–512 (1980).
83. Gold, T. *Proc. Roy. Soc. B135*, 492–498 (1948).
84. Weiss, T. F. *Hearing Res.* **7**, 353–360 (1982).
85. Neely, S. T. & Kim, D. O. *Hearing Res.* **9**, 123–130 (1983).
86. Zurek, P. M. *J. Acoust. Soc. Am.* **69**, 514–523 (1981).
87. Kemp, D. T. *J. Acoust. Soc. Am.* **64**, 1386–1391 (1978).
88. Kössel, M. & Vater, M. *Hearing Res.* **19**, 157–170 (1985).
89. Bialek, W. & Wit, H. P. *Phys. Lett.* **104A**, 173–178 (1984).
90. Spoendlin, H. in *Facts and Models in Hearing* (eds Zwicker, E. & Terhardt, E.) 18–32 (Springer, Berlin, 1974).
91. Brown, M. C. & Nuttall, A. L. *J. Physiol., Lond.* **354**, 625–646 (1984).
92. Dallos, P. & Harris, D. J. *Neurophysiol.* **41**, 365–383 (1978).
93. Brownell, W. E., Bader, C. R., Bertrand, D. & de Ribaupierre, Y. *Science* **227**, 194–196 (1985).
94. Kachar, B., Brownell, W. E., Altschuler, R. & Fex, J. *Nature* **322**, 365–368 (1986).
95. Ashmore, J. F. *J. Physiol., Lond.* **388**, 323–347 (1987).
96. Santos-Sacchi, J. & Dilger, J. P. *Hearing Res.* **35**, 143–150 (1988).
97. Ashmore, J. F. in *Cochlear Mechanisms: Structure, Function and Models* (eds Wilson, J. P. & Kemp, D. T.) 107–114 (Plenum, New York, 1989).
98. Holley, M. C. & Ashmore, J. F. *Nature* **335**, 635–637 (1988).
99. Zenner, H.-P., Zimmermann, R. & Gitter, A. H. *Hearing Res.* **34**, 233–240 (1988).
100. Dallos, P. *J. Neurosci.* **6**, 1591–1608 (1985).
101. Santos-Sacchi, J. *J. Neurosci.* **9**, 2954–2962 (1989).
102. Manley, G. A., Schulze, M. & Oelkinghaus, H. *Hearing Res.* **26**, 257–266 (1987).
103. van Dijk, P. & Wit, H. P. *J. Acoust. Soc. Am.* **82**, 2147–2150 (1987).
104. Wit, H. P., van Dijk, P. & Segenhout, J. M. in *Cochlear Mechanisms: Structure, Function and Models* (eds Wilson, J. P. & Kemp, D. T.) 341–347 (Plenum, New York, 1989).
105. Stryer, L. *Rev. Neurosci.* **9**, 87–119 (1986).
106. Lancet, D. A. *Rev. Neurosci.* **9**, 329–355 (1986).
107. Kuhn, G. F. in *Localization of Sound: Theory and Applications* (ed. Gatehouse, R. W.) 51–64 (Amphora, Groton, 1982).
108. Zwillocki, J. & Feldman, R. S. *J. Acoust. Soc. Am.* **28**, 860–864 (1956).
109. Moiseff, A. & Konishi, M. *J. Neurosci.* **1**, 40–48 (1981).
110. Gurd, F. R. N. & Rothgeb, T. M. *Adv. Protein Chem.* **33**, 73–165 (1979).
111. Catterall, W. A. *Science* **242**, 50–61 (1988).
112. Stimers, J. R., Bezanilla, F. & Taylor, R. E. *J. gen. Physiol.* **89**, 521–540 (1987).
113. Hudspeth, A. J. *Scient. Am.* **248**, 54–64 (1983).

ACKNOWLEDGEMENTS. The author thanks Drs P. G. Gillespie, A. H. Gitter, J. Howard, and L. F. A. Jaramillo for their comments on the manuscript, and colleagues at the University of California, San Francisco for a stimulating environment. Mr R. A. Jacobs fabricated the experimental apparatus, provided the electron micrographs of Figs 2 and 6, produced the figures and critically perused the manuscript. The NIH and the System Development Foundation supported the author's experimentation and the writing of this review.

Segmental expression of Hox-2 homoeobox-containing genes in the developing mouse hindbrain

David G. Wilkinson, Sangita Bhatt, Martyn Cook, Edorado Boncinelli* & Robb Krumlauf

Laboratory of Eukaryotic Molecular Genetics, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

* International Institute of Genetics and Biophysics, Via Guglielmo, Marconi 10, 80125 Naples, Italy

The vertebrate hindbrain develops in a segmental pattern, with distinctive groups of neurons originating from different segments. We report here that members of the Hox-2 cluster of murine homoeobox genes are expressed in segment-specific patterns in the developing hindbrain, with successive genes having boundaries at two-segment intervals. These data indicate that *Hox* genes specify segment phenotype, a role analogous to that of their *Drosophila* homologues.

A WIDESPREAD phenomenon in animal development is the process of segmentation, in which reiterated blocks of tissue generate homologous, yet distinct structures along the body axis. In higher vertebrates, it seems that in the trunk the mesodermal somites are the dominant segmental structures, whereas the central nervous system (spinal cord) is not intrinsically segmented¹. But the existence of repeated bulges (rhombomeres) during early stages of hindbrain development in vertebrates indicates that the hindbrain might be segmented²⁻⁸, and recent evidence has substantiated this view. Studies of the chick embryo have shown that rhombomeres(r) 2-6 are bounded by sulci that restrict specific populations of developing motoneurons⁹. Neurogenesis is initiated in alternating rhombomeres (r2, r4 and r6), and the Vth, VIIth and IXth cranial nerves arise from successive pairs of rhombomeres (r2 and 3, r4 and 5, and r6 and 7, respectively), indicating the existence of mechanisms operating at two-segment intervals⁹. Further evidence for segmentation of the hindbrain was found by analysis of the expression of the zinc-finger-encoding gene *Krox-20* in the mouse embryo¹⁰. The expression of *Krox-20* was found to be restricted to two alternate rhombomeres (r3 and r5) in the 9.5-day-old embryo, with the boundaries of expression occurring at rhombomeric sulci. Two domains of *Krox-20* expression were also detected in the hindbrain of the 8.5-day-old embryo, before the appearance of rhombomeres, indicating that segments are established at this stage. Although these data reveal the segmental mechanisms that underlie the organization of the hindbrain, the molecular basis of segmentation and of the specification of differences between the structures arising from each segment is not known.

In *Drosophila*, a diverse group of genes are involved in subdividing the embryo into segments, and the class of *Antennapedia*-like homoeobox genes (the HOM-C genes¹¹) specify segment phenotype according to position along the antero-posterior axis¹²⁻¹⁵. An important feature of this process is that these homoeobox genes are expressed in segment-restricted patterns. Murine homologues of the *Antennapedia*-class of homoeobox genes, termed the *Hox* genes, have been isolated¹⁶⁻²⁸ and shown to be organized in at least four clusters with homology to the *Drosophila* Bithorax (BX-C)²⁹ and An-

tennappedia (ANT-C)³⁰ complexes. Related genes in the murine and *Drosophila* clusters have extended sequence identity and are organized in the same order along the chromosome, indicating that they are derived from a common ancestor^{11,20,21}. Moreover, there is a similar correlation between the position of a gene in the *Drosophila* and mouse complexes and its pattern of expression along the antero-posterior axis of the embryo. Members at the 5' end of each cluster are expressed in posterior regions, and successively more 3' genes exhibit progressively more anterior domains of expression^{20,21,31}. In the mouse central nervous system (CNS), each of the *Hox* genes is expressed in a domain which extends from the caudal end of the neural tube to an anterior limit in the spinal cord or hindbrain³¹⁻⁴¹. In the absence of direct evidence for segmentation of the CNS, it has been suggested that these patterns indicate roles for *Hox* genes in positional specification independent of segmental processes. In light of the recent evidence for segmentation of the hindbrain, it seems that *Hox* genes could be involved in specifying segment phenotype in the neural tube. Previous studies have largely focused on stages subsequent to the disappearance of rhombomeres, so we examined the expression of Hox-2-cluster genes during the formation of hindbrain segments. We found that the boundaries of expression of the *Hox-2.6*, *-2.7* and *-2.8* genes are at the anterior sulci of r7, r5, and r3 in the 9.5-day-old embryo, and thus occur at two-segment intervals. By contrast, *Hox-2.9* expression is restricted to a single rhombomere, r4. Analysis of 8 to 8.5-day-old embryos showed that *Hox*-gene expression is initiated at the posterior end of the neural tube, and then extends to the appropriate anterior limit after the establishment of *Krox-20* expression, yet before the formation of overt rhombomeres. Segment-restricted *Hox* gene expression seems to occur therefore, after the formation of segments, consistent with *Hox* genes having a role in specifying segment phenotype rather than in the process of segmentation.

Rhombomeric expression of Hox-2-cluster genes

Earlier studies showed that the *Hox-2.1*, *-2.6* and *-2.7* genes are arranged in a 5'-3' order in the Hox-2 cluster and have progressively more anterior limits of expression in the hindbrain of the 12.5-day-old embryo^{19,20}. Two further genes, *Hox-2.8* and *-2.9*, have now been cloned and linked further downstream of the *Hox-2.7* gene in mice (M. Rubock, H. Lehrach, M.C. and R.K., unpublished results; M.C., R.K. and E.B., unpublished results) and humans^{42,52}. We analysed the expression of these genes in coronal sections of 9.5-day-old embryos, in which the presence of the otocyst adjacent to r5 and r6 (ref. 7) allows the identity of rhombomeres to be unambiguously assigned.

The boundary of *Hox-2.1* expression was located at approximately the hindbrain-spinal cord boundary (at the level of the 4th somite), and thus lies caudal to the most posterior hindbrain segment (Fig. 1a, b). The expression of *Hox-2.6* (Fig. 1c, d), *-2.7* (Fig. 1e-h) and *-2.8* (Fig. 1i-l), however, had anterior limits at the anterior sulci of r7 (the posterior sulcus of r6), r5 and r3, respectively. The boundaries of expression of these successive

Hox genes occur, therefore, at two-segment intervals, with the limit of *Hox-2.6* expression at the most posterior segment boundary. The expression of *Hox-2.7* and *-2.8* exhibited quantitative differences along the neural tube. Higher levels of *Hox-2.7* expression were detected in r5 than in more posterior segments (Fig. 1*e, f*), and *Hox-2.8* transcripts were more abundant in r3–r5 than in more caudal regions (Fig. 1*i, j*). In contrast to the other *Hox-2* genes, *Hox-2.9* expression was confined to a single rhombomere, r4, in the hindbrain (Fig. 1*m–p*), and is thus flanked by domains of *Krox-20* expression (Fig. 1*s, t*). In addition, *Hox-2.9* transcripts were detected more caudally in the spinal cord (data not shown). Expression of *Hox-2.8* (Fig. 1*k, l*) and *-2.9* (Fig. 1*q, r*) occurred in the VIIth/VIIIth ganglia (located adjacent to r4), whereas expression of *Hox-2.1*, *-2.6* or *-2.7* was not detected in these ganglia.

Establishment of segmental patterns

We have previously shown that *Krox-20* expression occurs in

two domains in the hindbrain of the 8.5-day-old embryo, and is later found restricted to r3 and r5 (ref. 10). These data indicate that segmentation of the hindbrain occurs before the appearance of rhombomeres, and that *Krox-20* expression can be used to mark the locations of prospective rhombomeric sulci. We analysed the onset of *Krox-20* expression in more detail, and found that initially one stripe of expression appears at 8 days of development (Fig. 2*a–d*), and that two domains of expression have formed at 8.5 days (Fig. 2*e, f*). The initial site of *Krox-20* expression is immediately adjacent to the preotic sulcus⁸ (pro-rhombomere A), which flanks the anterior stripe of *Krox-20* expression in the 8.5-day-old embryo, so we inferred that *Krox-20* expression occurs in the prospective r3 before it does in r5. At later stages the down regulation of expression also occurs in r3 before it does in r5 (Fig. 2*g, h*, and ref. 10). These findings may reflect the general antero-posterior direction of CNS maturation.

Expression of *Hox-2.6* (Fig. 3*a, b*), *Hox-2.7* (Fig. 3*d, e*), and

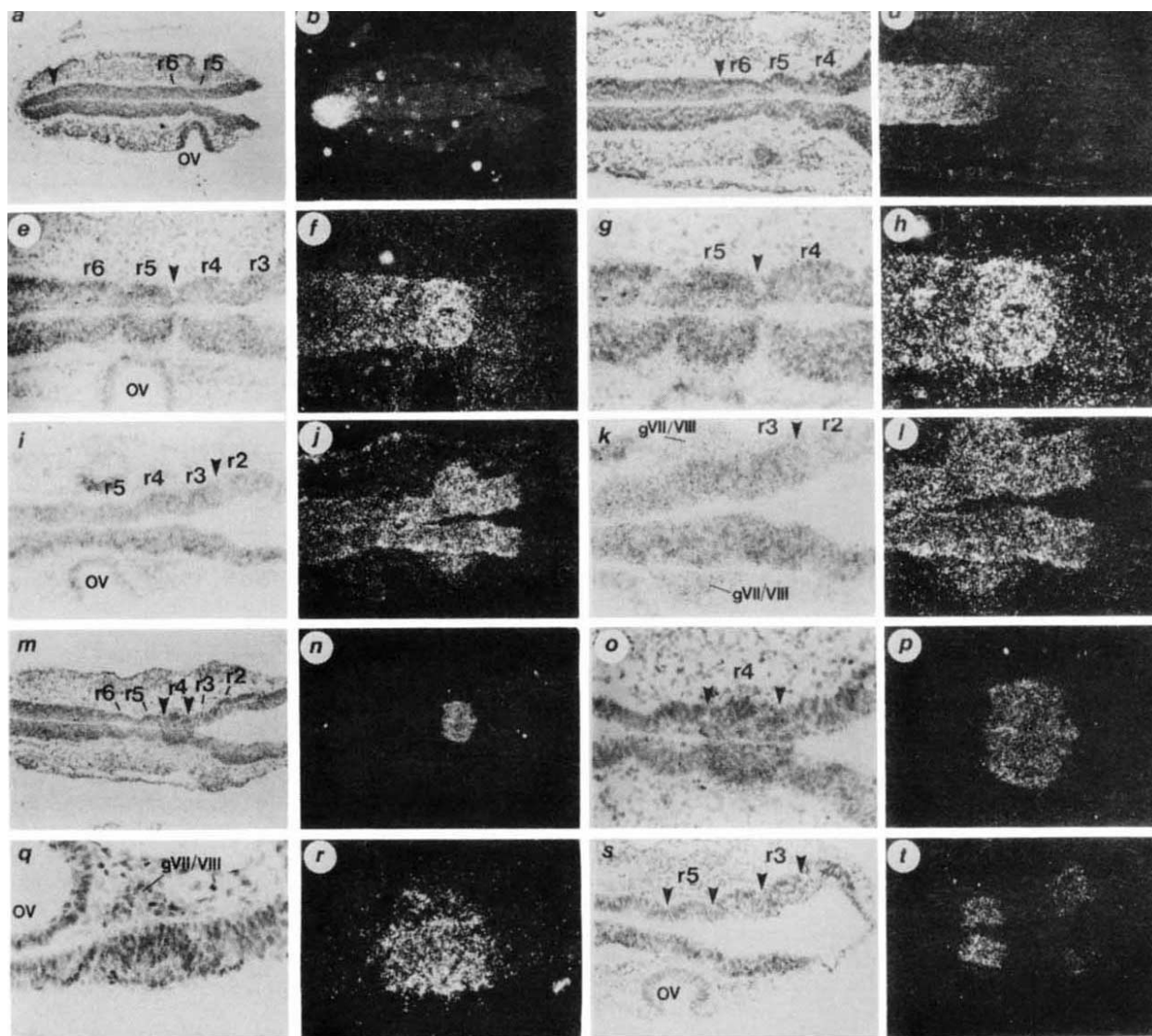


FIG. 1 Expression of *Hox-2*-cluster genes in the 9.5-day-old mouse embryo hindbrain. The relationship between *Hox-2*-gene expression and rhombomeres was determined by *in situ* hybridization of coronal sections of embryos with probes from the following genes: *a, b*, *Hox-2.1*; *c, d*, *Hox-2.6*; *e–h*, *Hox-2.7*; *i–l*, *Hox-2.8*; *m–r*, *Hox-2.9*; *s, t*, *Krox-20*. For each pair, the left micrograph is a brightfield photograph, and the right micrograph is the corresponding darkfield photograph. Rhombomeres are numbered and the boundaries of *Hox*-gene expression are indicated with arrows. OV, otic vesicle; gVII/VIII, VIIth/VIIIth ganglia. At this stage, rhombomeres are approx.

70 μ m in length. The micrographs were taken at the following magnifications: $\times 25$ (*a, b*); $\times 33$ (*c, d, s, t*); $\times 40$ (*i, j*); $\times 50$ (*e, f, m, n*); $\times 66$ (*k, l, o, p*); $\times 80$ (*g, h*); $\times 100$ (*q, r*).

METHODS. *In situ* hybridization of sections with ³⁵S-labelled RNA probes and washing at high stringency was performed as previously described⁵⁰. The probes corresponded to the following fragments: *Hox-2.1* and *-2.6* (ref. 20); *Hox-2.7*, 700 base pairs (bp) *Bam*HI–*Hind*III; *Hox-2.8*, 300-bp *Taq*I; *Hox-2.9*, 800-bp *Eco*RI; *Krox-20* (ref. 10).

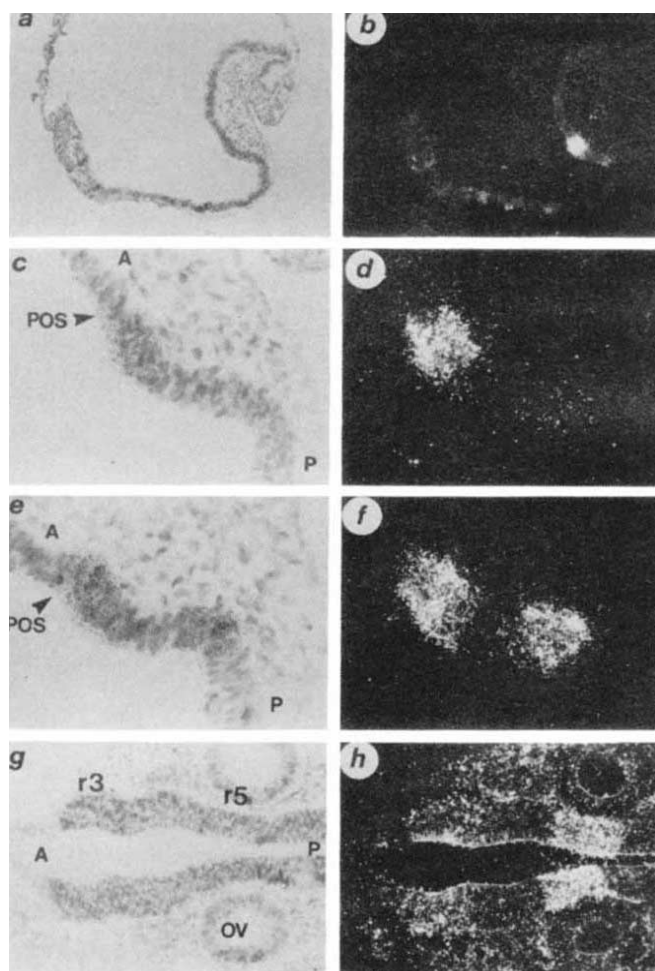


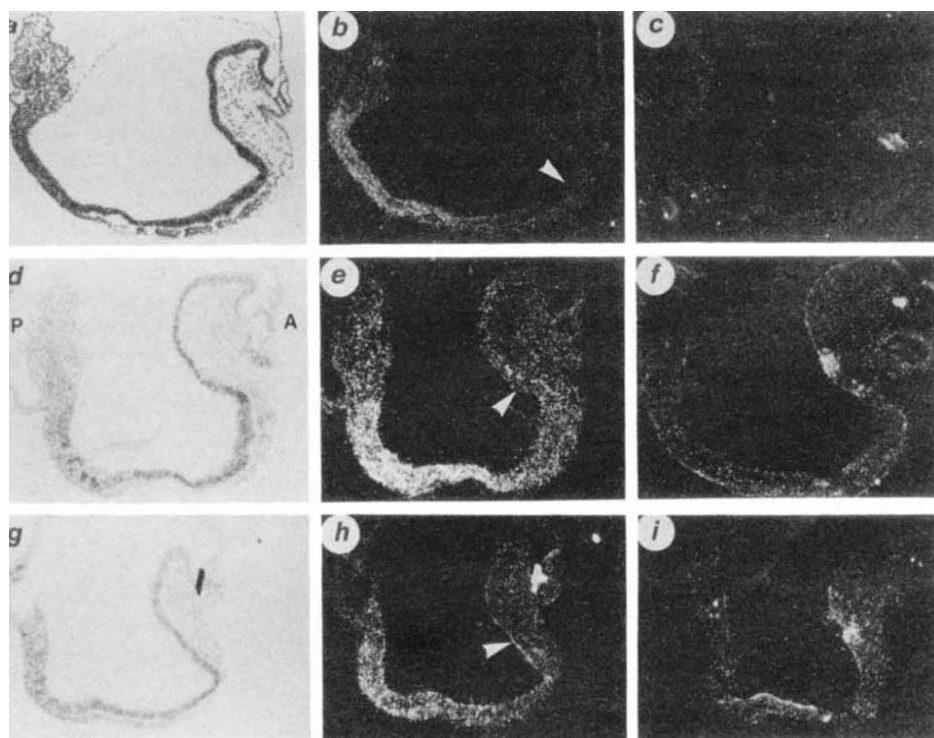
FIG. 2 Expression of *Krox-20* in 8 to 8.5-day-old mouse embryos. *In situ* hybridization of *Krox-20* was carried out against longitudinal (a–f) or coronal (g, h) sections of mouse embryos. a–d, 8.0-day-old embryo. e, f, 8.5-day-old embryo. g, h, 10-day-old embryo. Left column, brightfield photographs; each corresponds to its adjacent darkfield photograph. Rhombomeres are numbered. A, anterior; P, posterior; POS, preotic sulcus; OV, otic vesicle. Only low levels of *Krox-20* transcripts are detected in r3 in the 10-day-old embryo. Embryos were staged according to somite number as described by Theiler⁵¹. The micrographs were taken at the following magnifications: $\times 33$ (a, b); $\times 50$ (g, h); $\times 100$ (c–f). Experimental methods as described in the legend to Fig. 1.

Hox-2.9 (Fig. 3g, h) was detected only at the posterior end of the neural plate in 8-day-old embryos, which show only a single stripe of *Krox-20* expression. These domains of *Hox* expression are substantially caudal to the site of *Krox-20* expression (Fig. 3c, f, i) and to the anterior segmental limit of expression of these *Hox* genes found at 9.5-days of development (Fig. 1). But by 8.5 days *Hox-2.9* expression was detected in the hindbrain in a domain that is flanked by two stripes of *Krox-20* expression (Fig. 4a–f), and *Hox-2.8* expression extended to a boundary at the anterior-most limit of *Krox-20* expression (Fig. 4g–l). These positions of the *Hox-2.8* and *-2.9* expression domains relative to the *Krox-20* expression domain are the same as those in the 9.5-day-old embryo. Thus, it seems that the expression of *Hox* genes becomes established in a posterior–anterior manner, and that their segment-restricted domains of expression arise before the formation of definitive rhombomeres.

Segmental expression of *Hox* genes

Although the expression of different *Hox* genes in distinct overlapping domains is consistent with their having a role in specifying region-specific pattern, the biological significance of the anterior boundaries of expression was not apparent from previous studies. Based on studies of *Hox-1.5*-gene expression it has been suggested that the boundaries of *Hox*-gene expression might have a possible relationship with rhombomeres⁴¹. Here we have shown that the four 3'-most members of the *Hox-2* cluster are expressed in segment-restricted patterns, indicating that they play a part in specifying segment phenotype.

FIG. 3 Expression of *Hox-2*-cluster genes in 8-day-old mouse embryos. Longitudinal sections of embryos were hybridized with probes from the following genes: a, b, *Hox-2.6*; d, e, *Hox-2.7*; g, h, *Hox-2.9*; c, f, i, *Krox-20*. Left column, brightfield photographs; each corresponds to its adjacent darkfield photograph (centre columns). Centre and right columns, darkfield photographs of adjacent sections hybridized with *Hox-2*-gene and *Krox-20* probes, respectively. The location of the prospective boundaries of *Hox-2*-gene expression at 9.5-days (relative to *Krox-20* expression; see Fig. 1) are indicated with arrows. The prospective boundary of *Hox-2.6* indicated is immediately anterior to the first somite, as at 9.5 days of development. A, anterior; P, posterior. All micrographs were taken at $\times 33$ magnification. Experiments performed as described in the legend to Fig. 1.



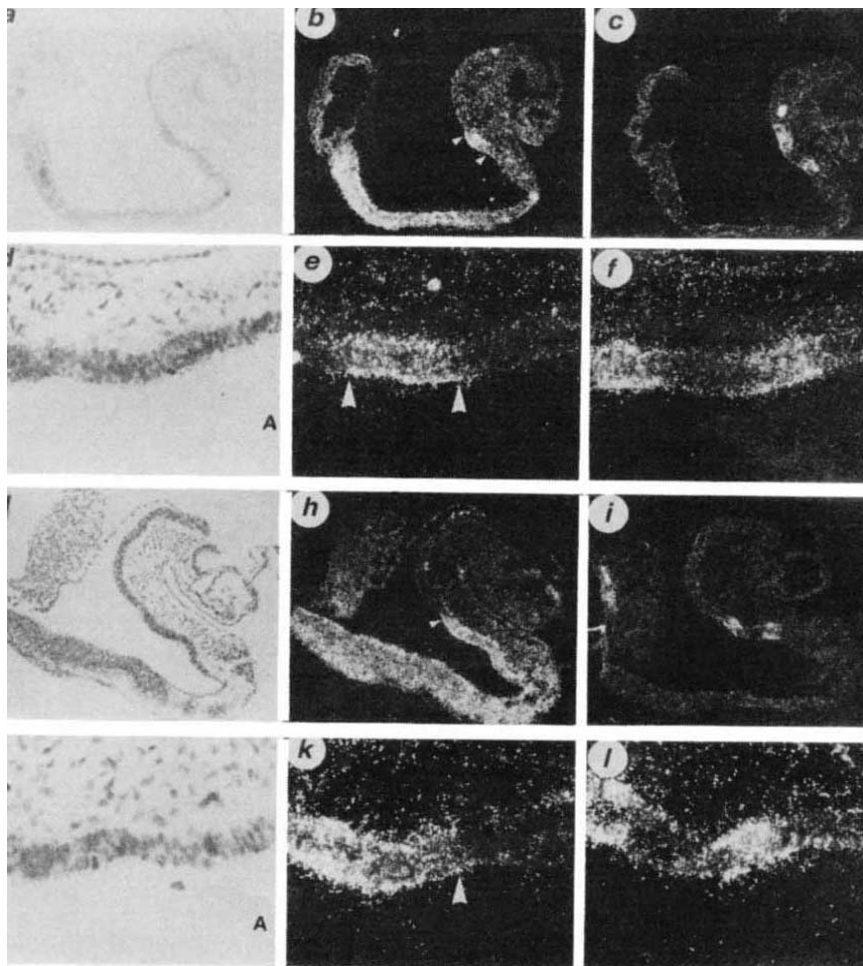


FIG. 4 Expression of Hox-2-cluster genes in 8.5-day-old embryos. Longitudinal sections of embryos were hybridized with the probes from the following genes: *a, b, d, e*, *Hox-2.9*; *g, h, j, k*, *Hox-2.8*; *c, f, i, l*, *Krox-20*. Left column, brightfield photographs; each corresponds to its adjacent darkfield photograph of the centre column. Centre and right columns, darkfield photographs of adjacent sections hybridized with Hox-2-gene and *Krox-20* probes, respectively. The higher-magnification photographs shown in *d-f*, and *j-l* are in the inverse orientation relative to those in *a-c*, and *j-l*. The locations of the limits of Hox-gene expression are at the same positions relative to the domains of *Krox-20* expression as they are 9.5 days. A, anterior; P, posterior. The micrographs were taken at the following magnifications: $\times 25$ (*a-c*); $\times 33$ (*g-i*); $\times 100$ (*d-f, j-l*). Experimental methods as described in the legend to Fig. 1.

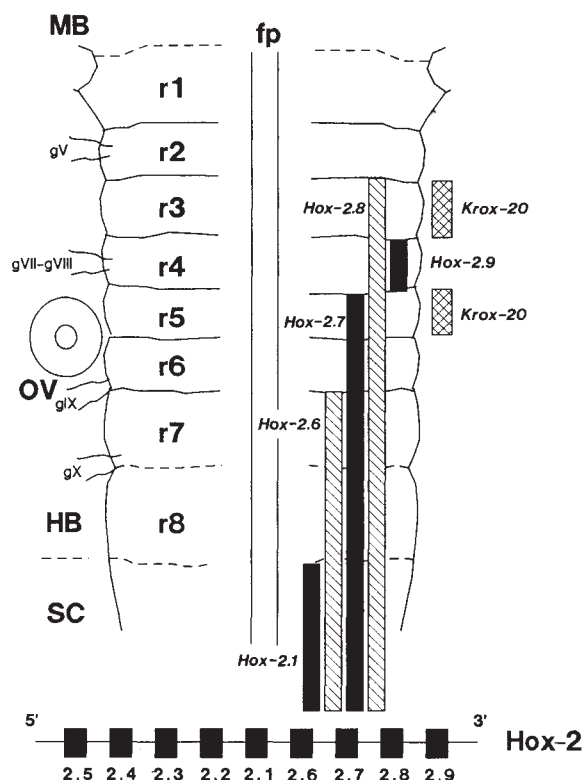


FIG. 5 Summary of Hox-2-cluster gene expression in the 9.5-day-old mouse hindbrain. The lower part illustrates the physical order of the Hox-2-cluster genes on chromosome 11. In the upper part, rhombomeres (r1-r8), the otic vesicle (OV) and the cranial ganglia (gV, gVII/VIII, gIX and gX) are represented on the left and the domains of Hox-2-gene and *Krox-20* expression are shown on the right. The anterior limits of r1 and r8 are indicated with dashed lines as they have not been shown to represent segment boundaries. MB, midbrain; HB, hindbrain; SC, spinal cord; fp, floor plate. For simplicity, the relationship between branchiomotor nerves and rhombomeres is not indicated; see Fig. 4 of Lumsden and Keynes (ref. 9) for details.

Our data argue against a function for these genes in the process of segmentation, because the anterior boundaries of *Hox*-gene expression are established after localized *Krox-20* expression. It seems that *Hox*-gene expression is established in a posterior-anterior direction along the neural tube until it reaches the appropriate previously-formed segment boundary. It seems, therefore, that the mechanisms that underlie the establishment of the domains of *Hox*-gene expression do not correlate with CNS maturation, which occurs in an anterior-posterior direction.

It is intriguing that the limits of *Hox-2.6*, *-2.7* and *-2.8* expression occur at two-segment intervals (Fig. 5), because the same periodicity has been shown by studies of the patterns of neuronal development⁹ and by the expression of *Krox-20*¹⁰. Pairs of rhombomeres express different combinations of these genes, although these are offset by one segment from the pairs of rhombomeres (r2 and 3, r4 and 5, r6 and 7) that contribute to individual branchial motor nerves (compare Fig. 5 with Fig. 4 of Lumsden and Keynes⁹). However, the VIth nerve, a somatic motor nerve, arises from r5 and 6, and this pair of rhombomeres lies in phase with the pattern of *Hox-2*-gene expression. It will be interesting to ascertain whether other hindbrain nuclei arise from pairs of rhombomeres offset from the branchiomotor system. The expression of *Hox-2.9* is restricted to r4 in the hindbrain and thus could be involved in specifying the phenotype of this particular segment. It is puzzling that expression of this gene does not continue the trend of increasingly anterior expression of successive *Hox-2* genes²⁰. We note, however, that continuation of the two-segment intervals between the boundaries of *Hox-2*-gene expression would predict a limit of *Hox-2.9* expression at the anterior end of r1, and that evidence that this rhombomere is a true segment is lacking. The patterns of expression that we have described are not sufficient to specify all the features of the hindbrain, in particular the pairs of rhombomeres that generate the branchiomotor system. Thus, it is pertinent to examine whether members of *Hox* clusters on other chromosomes exhibit segmental patterns of expression identical to, or offset from those of the *Hox-2* cluster³³.

Hox-2-cluster gene expression in the neural crest

Hindbrain segments have a number of functional interconnections with neural crest derivatives; for example, cranial ganglia

are associated with specific rhombomeres⁷, and the segmentally-arranged branchiomotor nerves innervate specific branchial arches⁹. A possible mechanism that might underlie these spatial relationships is that neural crest is patterned according to its rhombomeric origin^{9,43-45}. Therefore, it may be significant that only *Hox-2.8* and *-2.9* are expressed in the VIIth/VIIIth ganglia, located adjacent to r4, as it is only these *Hox-2* genes that are expressed in r4 (Fig. 5). Other *Hox-2* genes also exhibit restricted expression in the cranial ganglia, because *Hox-2.7* expression occurs in the IXth and Xth ganglia (A. Graham and R.K., unpublished results), and *Hox-2.1* expression occurs only in the Xth ganglion³³.

Evolution of segmental homoeobox gene expression

In this study we have focused on the *Hox-2*-cluster genes expressed in the segmented hindbrain. The significance of the boundaries of the other *Hox-2* genes (*Hox-2.5*, *-2.4*, *-2.3*, *-2.2* and *-2.1*) which lie in the spinal cord is not clear, however. For *Hox-2.1*, the limit of expression occurs in the vicinity of the hindbrain/spinal cord border, and thus may correlate with a transition in CNS organization. Although current evidence argues against the existence of segment boundaries in the caudal hindbrain and spinal cord of higher vertebrates, lower vertebrates do have segmentally arranged neurons in these regions^{46,47}. Segmentation seems, therefore, to have been retained in the upper hindbrain of higher vertebrates, but lost from more caudal parts of the neural tube⁹. It is possible that *Hox*-gene expression does respect spinal cord segments in lower vertebrates, and the location of these boundaries have been retained in higher vertebrates despite the loss of intrinsic segmentation.

The similarities in the structure and expression of *Antennapedia*-like homoeobox genes in *Drosophila* and mice indicate that they have an ancient conserved role in specifying regional variations in pattern along the body axis^{11,20,21}. Although such a function need not be coupled to segmentation, the finding that the expression of the *Hox* and HOM-C genes is segment-restricted in these diverse organisms is evidence for their having a conserved role in specifying segment phenotype. But because segmentation seems to have arisen separately in the lineages that lead to arthropods and vertebrates^{48,49}, the coupling of homoeobox-gene expression to segments may have evolved independently in mice and *Drosophila*. □

Received 8 August; accepted 1 September 1989.

1. Keynes, R. J. & Stern, C. D. *Development* **103**, 413-431 (1988).
2. Orr, H. J. *Morph.* **1**, 311-372 (1987).
3. Streeter, G. L. *Anat. Rec.* **2**, 111-115 (1908).
4. Neal, H. V. J. *Morph.* **31**, 293-315 (1918).
5. Adelmann, H. B. J. *comp. Neurol.* **39**, 19-171 (1925).
6. Berquist, H. & Kallen, B. J. *comp. Neurol.* **100**, 627-660 (1954).
7. Vaage, S. *Adv. Anat. Embryol. Cell Biol.* **41**, 1-88 (1969).
8. Tuckett, F., Lim, L. & Morriss-Kay, G. M. J. *Embryol. exp. Morph.* **87**, 315-228 (1985).
9. Lumsden, A. & Keynes, R. *Nature* **337**, 424-428 (1989).
10. Wilkinson, D. G., Bhatt, S., Chavrier, P., Bravo, R. & Charnay, P. *Nature* **337**, 461-464 (1989).
11. Akam, M. *Cell* **57**, 347-349 (1989).
12. Gehring, W. J. *Science* **236**, 1245-1252 (1987).
13. Akam, M. *Development* **101**, 1-22 (1987).
14. Akam, M., Dawson, I. & Tear, G. *Development* **104** suppl. 123-133 (1988).
15. Ingham, P. W. *Nature* **335**, 25-34 (1988).
16. Duboule, D., Baron, A., Mahi, P. & Galliot, F. *EMBO J.* **5**, 1973-1980 (1986).
17. Colberg-Poley, A. M. et al. *Cell* **43**, 39-45 (1985).
18. Papalopulu, N. et al. in *Cell to Cell Signals in Mammalian Development*, NATO ASI series **H26**, 9-21 (Springer, Berlin, 1989).
19. Graham, A. et al. *Genes Dev.* **2**, 1424-1438 (1988).
20. Graham, A., Papalopulu, N. & Krumlauf, R. *Cell* **57**, 367-378 (1989).
21. Duboule, D. & Dolle, P. *EMBO J.* **8**, 1497-1505 (1989).
22. Hart, C. P., Awgulewitsch, A., Fainsod, A., McGinnis, W. & Ruddle, F. H. *Cell* **43**, 9018 (1985).
23. Hart, C. P., Fainsod, A. & Ruddle, F. H. *Genomics* **1**, 182-185 (1987).
24. Lonai, P., Arman, E., Czosnek, H., Ruddle, F. H. & Blatt, C. *DNA* **6**, 409-418 (1987).
25. Breier, G., Dressler, G. R. & Gruss, P. *EMBO J.* **7**, 1329-1336 (1988).
26. Sharpe, P. T., Miller, J. R., Evans, E. P., Burtenshaw, M. D. & Gaunt, S. J. *Development* **102**, 397-407 (1988).
27. Schughart, K. et al. *Genomics* (in the press).
28. Featherstone, M. S., Baron, A., Gaunt, S. J., Mattei, M. & Duboule, D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4760-4764 (1988).
29. Kaufman, T. C., Lewis, R. & Wakimoto, B. *Genetics* **94**, 115-133 (1980).

30. Lewis, E. B. *Nature* **276**, 565-570 (1978).
31. Gaunt, S. J., Sharpe, P. T. & Duboule, D. *Development* **104** suppl. 169-179 (1988).
32. Krumlauf, R., Holland, P. W. H., McVey, J. H. & Hogan, B. L. M. *Development* **99**, 603-617 (1987).
33. Holland, P. W. H. & Hogan, B. L. M. *Development* **102**, 159-174 (1988).
34. Dony, C. & Gruss, P. *EMBO J.* **6**, 2965-2975 (1987).
35. Fibi, M. et al. *Development* **102**, 349-360 (1988).
36. Schughart, K., Utset, M. F., Awgulewitsch, A. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5582-5586 (1988).
37. Awgulewitsch, A., Utset, M. F., Hart, C. P., McGinnis, W. & Ruddle, F. H. *Nature* **320**, 328-335 (1986).
38. Toth, L. E., Slawin, K. L., Pintar, J. E. & Chi Nguyen-Huu, M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6790-6794 (1987).
39. Gaunt, S. J., Miller, J. R., Powell, D. J. & Duboule, D. *Nature* **324**, 662-664 (1986).
40. Gaunt, S. J. *Development* **103**, 135-144 (1988).
41. Gaunt, S. J. *Development* **101**, 51-60 (1987).
42. Boncinelli, E. et al. *Hum. Reprod.* **3**, 880-886 (1988).
43. Noden, D. M. *Dev. Biol.* **96**, 144-165 (1983).
44. Noden, D. M. *Development* **103** suppl. 121-140 (1988).
45. Thorogood, P. in *Mesenchymal-Epithelial Interactions in Neural Development*, 141-152 (Springer, New York, 1987).
46. Kimmel, C. B., Sepich, D. S. & Trevarrow, B. *Development* **104**, suppl. 197-207 (1988).
47. Hanneman, E., Trevarrow, B., Metcalfe, W. K., Kimmel, C. & Westerfield, M. *Development* **103**, 49-58 (1988).
48. Hyman, L. H. *The Invertebrates: Platyhelminthes and Rhynchocoela* Vol II (McGraw-Hill, New York, 1951).
49. Clark, R. B. *Dynamics in Metazoan Evolution* (Oxford University Press, 1964).
50. Wilkinson, D. G., Bailes, J. A. & McMahon, A. P. *Cell* **50**, 79-88 (1987).
51. Theiler, K. *The House Mouse* (Springer, Berlin, 1972).
52. Boncinelli, E. et al. *Genome* **31**, 128-143 (1989).
53. Gaunt, S. J., Krumlauf, R. & Duboule, D. *Development* **107**, 131-142 (1989).

ACKNOWLEDGEMENTS. We thank Dario Acampora, Anthony Graham and Andrew Lumsden for suggestions and discussions. We also thank Melissa Rubock and Hans Lehrach for allowing us to refer to the cloning and characterization of the mouse *Hox-2.8* gene prior to publication. Part of the work was supported by CNR Project on Biotechnology and by AIRC.

Structure and associated DNA-helicase activity of a general transcription initiation factor that binds to RNA polymerase II

Mary Sopta, Zachary F. Burton* & Jack Greenblatt

Department of Medical Genetics & Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario M5G 1L6, Canada

* Department of Biochemistry, Michigan State University, East Lansing, Michigan 48824, USA

RAP30/74 is a heteromeric general transcription initiation factor which binds to RNA polymerase II. Here we report that preparations of RAP30/74 contain an ATP-dependent DNA helicase whose probable function is to melt the DNA at transcriptional start sites. The sequence of the RAP30 subunit of RAP30/74 indicates that RAP30 may be distantly related to bacterial sigma factors.

ACCURATE initiation of transcription by RNA polymerase II requires several auxiliary protein factors and ATP¹. The best characterized transcription factors are the sequence-specific DNA-binding proteins, including promoter-binding proteins such as TFIID (the TATA-binding factor)²⁻⁶ and Sp1 (ref. 7), as well as proteins that bind to enhancers and activate transcription in response to particular stimuli. Initiation by RNA polymerase II also requires several less well characterized general initiation factors, sometimes referred to as TFIIB, TFIIE and TFIIF⁸⁻¹⁰, which apparently do not recognize DNA in a sequence-specific manner. The general initiation factors are known to interact with RNA polymerase II¹¹⁻¹³, or with each other¹², and enter the initiation complex late in the initiation pathway¹²⁻¹⁵. Their roles in transcription have remained obscure.

We previously identified and purified one of the general initiation factors by affinity chromatography on a column containing immobilized RNA polymerase II^{11,15}. This factor, which we call RAP30/74, has two subunits^{11,16} with apparent relative molecular masses of 30,000 and 74,000 (RAP30 and RAP74 subunits, respectively). RAP30/74 is required for formation of the first phosphodiester bond at a promoter^{15,16}. It is needed for initiation, regardless of whether or not the promoter has a TATA box¹⁶. RAP30/74 is similar or identical to TFIIF¹⁰, which contains at least the RAP30 subunit of RAP30/74 (ref. 10). TFIIF is one of the components of a less purified fraction, called TFIIE, which has been widely used in many studies of initiation by RNA polymerase II^{17,18}. General initiation factors that may be equivalent to RAP30/74 or to its RAP30 subunit have also been described^{13,19}.

The ability to purify RAP30/74 by affinity chromatography¹¹ has now enabled us to clone human complementary DNAs that encode RAP30 and to begin the biochemical characterization of RAP30/74. Here we report that RAP30/74 has an associated DNA-helicase activity whose requirement for ATP may partly explain the energy requirement for transcription initiation by RNA polymerase II¹.

Associated DNA-helicase activity

We considered the possibility that RAP30/74 might be required for initiation by RNA polymerase II because it is required to melt the DNA template at the start site of transcription. Accordingly, we tested whether our preparations of RAP30/74 had DNA-melting (that is, helicase) activity.

We purified RAP30/74 from a HeLa-cell nuclear extract by

affinity chromatography on a column containing immobilized RNA polymerase II^{11,15,16}. To be certain that any activity eluting from the column was associated with RAP30/74, we performed the purification in duplicate (Fig. 1a). We pre-treated the HeLa-cell nuclear extract either with antibodies directed against a TrpE-RAP30 fusion protein to deplete the extract of RAP30/74 (Fig. 1b, lane 2), or with antibody directed against TrpE (Fig. 1b, lane 3) before we loaded it onto either an RNA polymerase II column or a control column. We eluted these columns (Fig. 1a) with salt, so that RAP30/74 dissociated from RNA polymerase II¹¹. These high-salt eluates were then tested for DNA-helicase activity in a standard assay^{20,21} requiring the displacement of a small radiolabelled complementary oligonucleotide which had been annealed to single-stranded M13 DNA (Fig. 1c).

The high-salt eluate of the RNA polymerase II column that had been loaded with nuclear extract treated with anti-TrpE antibody had DNA-helicase activity (Fig. 1c, lane 6). We confirmed that this was the only eluate that contained RAP30 by western blotting with anti-RAP30 antibody (Fig. 1b, lane 6). The dose-response curve for the DNA-helicase activity was roughly linear (Fig. 1d) and no unwinding was observed after about 5 minutes (data not shown). The DNA-helicase activity of RAP30/74 was not affected by NaCl concentrations as high as 200 mM (data not shown), even though some DNA helicases are inactive or strongly inhibited under these conditions²².

The DNA-helicase activity was able to bind to an RNA polymerase II column and not to a control column (see lanes 6 and 7 in Fig. 1c, and also Fig. 1d); it was specifically removed from a nuclear extract by affinity-purified antibody directed against RAP30 (compare lanes 4 and 6 in Fig. 1c). The two subunits of RAP30/74 are the only general initiation factors that bind efficiently to RNA polymerase II^{11,15}. Because they are the only proteins in a HeLa-cell nuclear extract that are bound efficiently by both anti-RAP30 antibody and RNA polymerase II^{11,15,16}, either RAP30/74 is a DNA helicase or the DNA helicase is a different protein that binds to a complex containing both RAP30/74 and RNA polymerase II.

Because the DNA-helicase activity associated with RAP30/74 was not directly inhibited by affinity-purified polyclonal antibody directed against RAP30 (data not shown), it is unlikely that the RAP30 subunit of RAP30/74 is necessary for DNA helicase activity. We tested whether RAP30 or RAP74 could bind to ATP, because DNA helicases must bind DNA and generally require ATP for their activity²³. We found that RAP74 can form a covalent complex containing the alpha-phosphate from ATP, and that the formation of this complex is stimulated by DNA (Fig. 2). These observations are consistent with the idea that the RAP74 subunit of RAP30/74 is the DNA helicase.

Nucleotide hydrolysis requirements

RNA polymerase II can accurately initiate transcription only if hydrolysable ATP or dATP is present in the reaction^{1,24}. We therefore investigated whether the requirement for ATP in the initiation of transcription reflects a requirement of RAP30/74 for energy in melting a DNA duplex. The DNA-helicase activity

of our RAP30/74 preparation required ATP and could use dATP (lanes 3 and 5, Fig. 3a), but could not use GTP, CTP, UTP, dGTP, dCTP or TTP (lanes 6–11). Half-maximal activity was achieved at an ATP concentration of 150 μ M (Fig. 3b). ATP hydrolysis was necessary because the ATP analogue, AMP-PNP (β - γ -Imidoadenosine-5'-triphosphate), could not substitute for ATP (Fig. 3a, lane 4). These features of the DNA-helicase activity associated with RAP30/74 are exactly the same as those that are characteristic of the energy requirement for initiation by RNA polymerase II^{1,24}.

Sarkosyl sensitivity

Initiation of transcription by a complete rapid-start initiation complex at the adenovirus major late promoter is inhibited by Sarkosyl detergent²⁵. The DNA-helicase activity associated with RAP30/74 was inhibited by the same concentrations of Sarkosyl that inhibit initiation by a rapid-start complex²⁵ (Fig. 3c). This indicates that the Sarkosyl sensitivity of a transcriptional initiation complex reflects the Sarkosyl sensitivity of a DNA helicase required to melt the DNA duplex at the promoter.

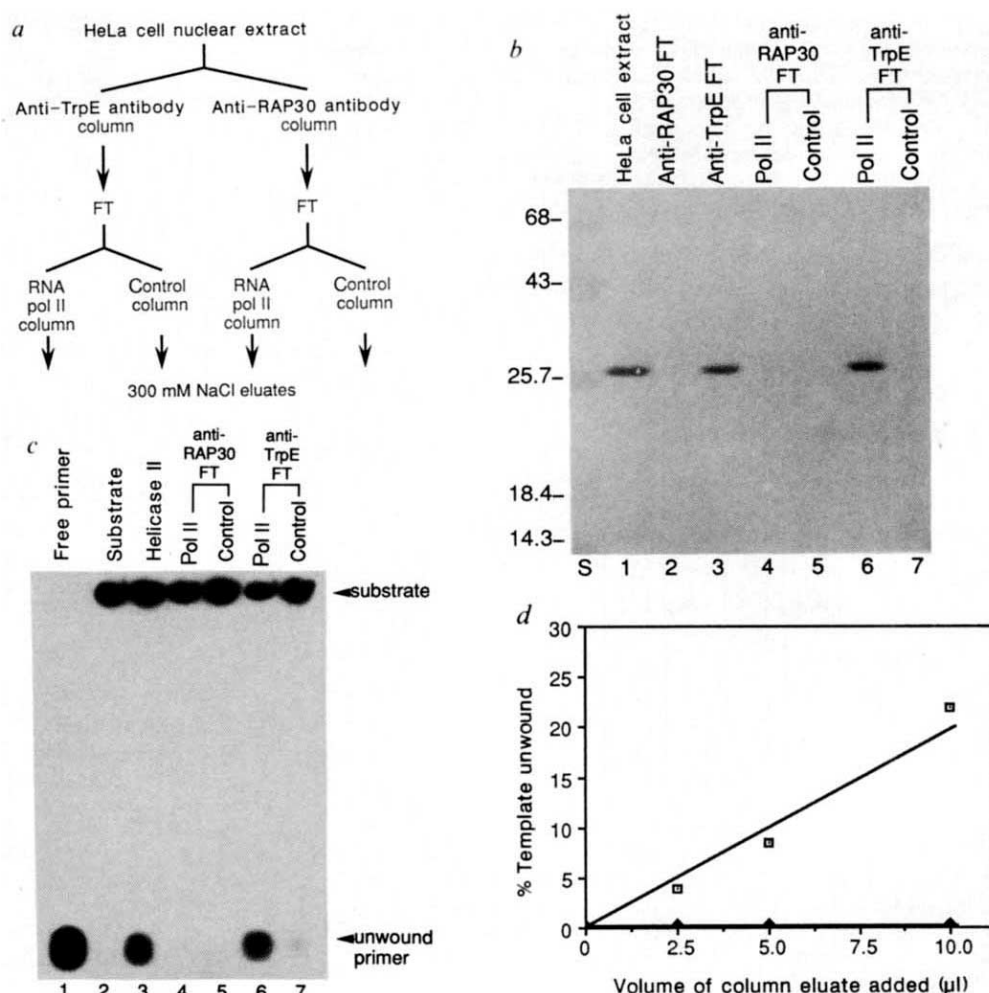
Sequence of RAP30 subunit of RAP30/74

To analyse the structure and function of RAP30/74, we cloned and sequenced cDNAs encoding the RAP30 subunit. We purified human RAP30 from HeLa cells by RNA polymerase II-affinity chromatography and SDS-polyacrylamide gel electrophoresis, and used it to immunize a rabbit. With the resulting antiserum we screened an expression library in λ gt11 (ref. 26) made with cDNAs derived from the human promyelocytic-leukaemia cell line, HL60. A partial RAP30 cDNA obtained in this way was then used as a hybridization probe to isolate cDNAs containing the entire RAP30 coding sequence from a human kidney cDNA library. The 1,426-base pair (bp) DNA sequence up to the first polyadenylation site is shown in Fig. 4. The open reading frame begins with ATG in a context that is typical of mammalian translational start sites²⁷. It encodes a protein of 234 amino acids with a predicted relative molecular mass of 26,312.

We confirmed that our cloned cDNAs encoded RAP30. First, affinity-purified antibodies against RAP30- β -galactosidase and TrpE-RAP30 fusion proteins produced in *Escherichia coli*

FIG. 1 RAP30 has an associated DNA-helicase activity. *a*, Scheme for the preparation of column fractions. *b*, Western-blot analysis of column fractions using anti-RAP30 antibody. *c* and *d*, DNA helicase activities in RNA polymerase II and control column eluates. Polymerase II denoted by pol II. FT, Flow through.

METHODS. *a*, HeLa-cell nuclear extract (1 ml)⁴⁶, containing ~ 4 ng μ l⁻¹ of RAP30, was passed through 0.5-ml columns containing immobilized affinity-purified anti-RAP30 antibodies or anti-TrpE antibodies (see below) at a concentration of 360 μ g ml⁻¹. The flow-through from each was then subjected to affinity chromatography on 0.5-ml columns containing immobilized RNA polymerase II (2 mg ml⁻¹ protein) and on control columns¹¹. The affinity columns were eluted with affinity column buffer¹¹ containing 0.3 M NaCl. *b*, The amounts loaded onto the gel were (μ l): HeLa-cell nuclear extract, 5; flow-throughs from anti-RAP30 and anti-TrpE antibody columns, 5; 0.3 M NaCl eluates of RNA polymerase II and control columns, 15. The blot was developed with an alkaline phosphatase-based second antibody reaction (BioRad). Pre-stained molecular weight markers (lane 5) were from BioRad. *c* and *d*, An M13 displacement reaction was used to assay DNA helicase activity^{20,21}. The substrate was an end-labelled universal sequencing primer hybridized to M13 single-stranded DNA. DNA helicase assays (20 μ l; 20 min at 37 °C) contained 40 mM Tris-HCl, pH 7.9, 4 mM MgCl₂, 150 mM NaCl, 1 mM dithiothreitol (DTT), 50 μ g ml⁻¹ BSA, 2 mM ATP, 0.2 nM DNA substrate, and 10 μ l (*c*) or the indicated volumes (*d*) of the column eluates. Reactions were electrophoresed on 15% non-denaturing polyacrylamide gels⁴⁷. The dried gels were exposed to X-ray film Kodak XAR-5. A reaction with 40 ng *E. coli* DNA helicase II (provided by Dr Steve Matson) contained only 75 mM NaCl. The reaction in lane 1 in *c* was incubated for 1 min at 100 °C. The percentage of DNA unwound (*d*) was calculated after subtraction of a background value of 5%. Preparation of anti-RAP30 and anti-TrpE antibodies: The TrpE-RAP30 fusion plasmid pATHR30 was constructed by cloning the RAP30-encoding *Pst*I fragment of p9-2 (see Fig. 3 legend) into the *Pst*I site of pATH10 (ref.



48). The strain JM101 containing either the plasmid pATH10 or pATHR30 was grown as previously described⁴⁸ (10 litres of each at 30 °C). TrpE and TrpE-RAP30 were purified from cleared sonicates by precipitation with $(\text{NH}_4)_2\text{SO}_4$ at 25% saturation. Yields were 75 mg of each at 80–90% purity. Rabbits were immunized¹⁶ with purified TrpE-RAP30. The resulting antisera were affinity-purified on columns containing immobilized TrpE to produce anti-TrpE antibodies (ref. 16). The flow-throughs were purified on columns containing immobilized TrpE-RAP30 to produce anti-RAP30 antibodies (ref. 16).

reacted with human RAP30 in a western blot¹⁶ (Fig. 1b). Such antibodies also efficiently and specifically removed RAP30 and RAP74 (by virtue of its tight interaction with RAP30) from a HeLa-cell nuclear extract¹⁶. Transcriptional activity could only be fully restored to such an extract by the addition of a combination of gel-purified RAP30 and gel-purified RAP74¹⁶. Second, we determined the N-terminal sequences of three tryptic fragments and one endoproteinase Asp-N fragment of RAP30 that had been purified from HeLa cells by RNA polymerase II-affinity chromatography and gel electrophoresis. These matched perfectly with the amino-acid sequences that were predicted from the RAP30 cDNA sequence (Fig. 4).

An ATP-dependent DNA helicase must bind to DNA and to ATP, and we have shown here that RAP74 has a DNA-regulated ATP-binding site (Fig. 2). We have no evidence that RAP30 can bind to ATP, but there is a Gly-X-X-X-X-Gly-Lys motif at residues 36–42 of RAP30, which is often a part of nucleotide-binding sites²⁸. We also examined the predicted amino-acid sequence of RAP30 for canonical 'helix-turn-helix' or 'zinc-finger' DNA-binding motifs^{29,30}, but there were no obvious candidate sequences in RAP30. RAP30 is very basic (pI = 10.4), with a region of dense positive charge at amino-acid residues 168–182 (Fig. 4). This stretch of amino acids is predicted to form an α -helical structure. The function of this region is unknown, although certain DNA-binding regions³¹ are positively charged α -helical structures. Interestingly, the sequence of RAP30 in this region matches 5/5 or 7/11 with half of the proposed DNA-binding region of CREB³², a DNA-binding protein which mediates responsiveness to peptide hormones that raise intracellular cyclic AMP (Fig. 5c). RAP30 has none of the sequence motifs often found in other DNA helicases³³.

RAP30 may belong to the σ -factor family

The predicted amino-acid sequence of RAP30 is not closely

related to any other amino-acid sequences in the GENBANK, EMBL and PIR databases. But because RAP30 binds to RNA polymerase II, we compared the sequence of RAP30 directly with those of two bacterial RNA polymerase-binding proteins, NusA and σ ⁷⁰. There was no discernible relationship between RAP30 and the *E. coli* elongation factor NusA (ref. 34), but RAP30 did contain a region of similarity with the *E. coli* initiation factor σ ⁷⁰ and its *Bacillus subtilis* equivalent, σ ⁴³

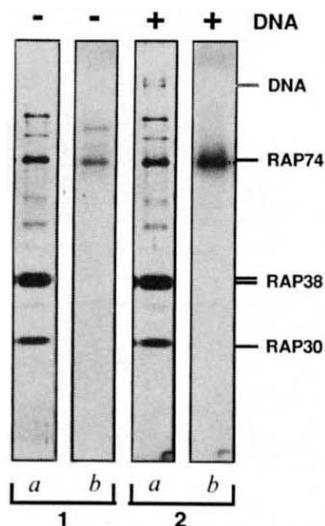


FIG. 2 Covalent labelling of RAP74 with [α -³²P] ATP. An RAP fraction radiolabelled with ATP in the presence or absence of DNA was subjected to SDS-PAGE. Lanes 1a and 2a, silver-stained RAP fractions. Lanes 1b and 2b, ³²P-autoradiography of the same gel lanes. The reaction in lane 2 contained DNA.

METHODS. RAP samples were isolated by affinity chromatography on an RNA polymerase II column and then by ion-exchange chromatography on BioRex 70 (ref. 16). RAP samples were incubated for 1 min on ice in 20 μ l buffer containing 12 mM HEPES buffer, pH 7.9, 12% glycerol, 13 mM KCl, 5 mM MgCl₂, 0.3 mM DTT, 0.12 mM EDTA, and 0.62 M [α -³²P] ATP (811 Ci mmol⁻¹) in the absence or presence of 20 μ g ml⁻¹ pSmaF DNA digested with *Sma*I (pSmaF is a plasmid containing the adenovirus major late promoter; it is not clear that promoter DNA is required for this labelling reaction). Samples were boiled in buffer containing 1% SDS and 1% 2-mercaptoethanol and electrophoresed. Radiolabelling of RAP74 was stimulated >4-fold by the addition of DNA to the reaction.

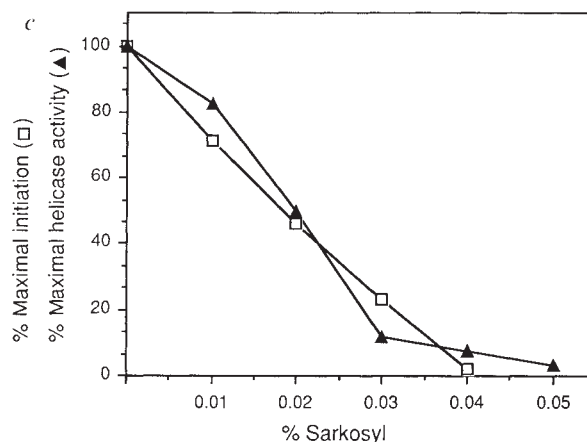
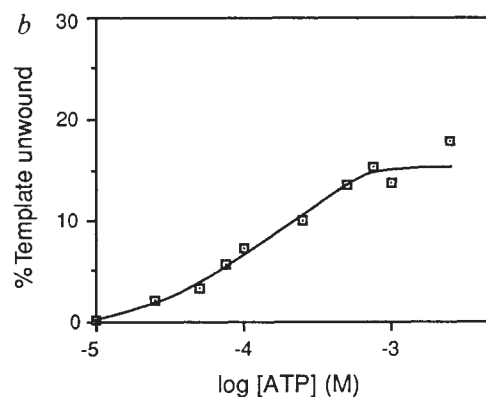
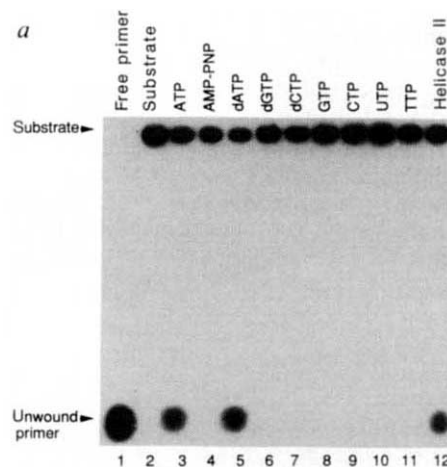


FIG. 3 Nucleotide requirement and Sarkosyl sensitivity of the DNA helicase. Assay conditions were as described in the legend to Fig. 1, except that 2 mM ATP was replaced with the equivalent concentrations of the indicated alternative nucleotides (a) or with the indicated concentrations of ATP [ATP] (b), and Sarkosyl was included at the indicated concentrations in c. The data in c on the effect of Sarkosyl on initiation by rapid-start complexes at the adenovirus major late promoter are taken from Hawley and Roeder²⁵. All nucleotides were purchased from Boehringer. *E. coli* DNA-helicase II (20 ng) was used as a standard.

(Fig. 5a). In particular, the best matches between RAP30 and σ^{70} were in regions 2 and 1b of σ^{70} , the only regions that are conserved among all bacterial and bacteriophage sigma factors. These parts of σ^{70} have been postulated to contain the site of the interaction of σ^{70} with the core component of bacterial RNA polymerase³⁵⁻³⁷. Indeed, a tryptic fragment of *E. coli* σ^{70} that retains the ability to bind to the core component of RNA polymerase³⁸ has end points that are nearly coincident with the boundaries of the proposed homology between σ^{70} and human RAP30 (Fig. 5a). An insertion of 245 amino acids in this region is found only in the sequence of *E. coli* σ^{70} . This insertion seems not to have an essential role in the interaction of σ^{70} with RNA polymerase, because it is missing in other σ -factors and because the *rpoD800* deletion in this sequence is not lethal³⁹.

We used the PIR ALIGN program for statistical analysis of the alignment between RAP30 and σ^{70} in homology region 2. The alignment score was marginally significant⁴⁰ and was comparable to, or better than, the alignment scores obtained for comparisons between σ^{70} and bacteriophage σ -factors in this region³⁶ (Fig. 5b), but low relative to those for comparisons among the main σ -factors of *E. coli* and *B. subtilis*. This indicates that, like the bacteriophage σ -factors, RAP30 may be a distantly related member of the σ -factor protein family.

Discussion

We have shown that the RAP30 subunit of the polymerase-binding general initiation factor, RAP30/74, may be distantly related to bacterial σ -factors. Indeed, the best sequence match between human RAP30 and *E. coli* σ^{70} lies in the only part of σ^{70} that is conserved among all σ -factors and is believed to be the RNA polymerase-binding region of σ^{70} (see refs 35 and 36). Bacterial σ -factors recognize promoter sequences and position the core component of RNA polymerase at transcriptional start sites⁴¹. In eukaryotes, the promoter-recognition function of the σ -factors seems to have become the responsibility of separate DNA-binding proteins such as the TATA-binding factor TFIID²⁻⁶ or Sp1 (ref. 7). By contrast, it seems that the RAP30

subunit of RAP30/74 has assumed the polymerase-binding function of bacterial σ -factors. We have shown previously that RAP30 is bound very stably to rapid-start initiation complexes at the adenovirus major late promoter^{10,16}, but we do not yet know whether, like σ^{70} (ref. 42), RAP30/74 is released from the transcription complex after the initiation of transcription.

We have also noted a possibly significant resemblance between the highly basic amino acids 164-174 of RAP30 and part of the DNA-binding site of CREB. This implies that RAP30 could be a site-specific DNA-binding protein with limited specificity. Because RAP30 seems not to be the TATA-binding factor^{10,15,16}, one possibility so far untested is that it could recognize the recently identified initiator element⁴³ located at some transcriptional initiation sites.

We have also shown that our preparations of RAP30/74 contain a DNA-helicase activity which must be bound, either directly or indirectly, to RAP30. The most obvious candidate for the DNA helicase is RAP74, but we have still not formally excluded the possibility that the DNA helicase is another polypeptide that binds to RAP30. We suggest that at least part of the role of RAP30 is to properly position the DNA helicase in the initiation complex by interacting with RNA polymerase II. The DNA-helicase would then melt the DNA at the correct site and cause isomerization from a closed complex to an open complex by analogy with the situation in bacteria⁴¹.

The distance from a TATA box to a transcriptional start site for mammalian promoters is different from that for promoters

FIG. 4 Sequence of human RAP30 cDNA. The boxed sequence represents the region of RAP30 that matches best with *E. coli* σ^{70} (Fig. 4). The major antigenic epitopes of human RAP30 are found in its 66 C-terminal amino acids (unpublished data). A potential nucleotide binding motif (Gly-X-X-X-Gly-Lys)²⁸ is found at amino-acid residues 36-42. The four amino-acid sequences underlined in bold indicate the N-terminal sequences that were determined for three tryptic fragments and one endoproteinase Asp-N fragment derived from human RAP30. The bar at the bottom shows the positions of various motifs in the RAP30 coding sequence and the distributions of positively and negatively charged amino acids in the protein.

METHODS. Rabbit antiserum raised against human RAP30 that was purified by RNA polymerase II-affinity chromatography and SDS-gel electrophoresis¹⁶ was used to screen a cDNA expression library in λ gt11²⁶ derived from HL60 human promyelocytic leukaemia cell messenger RNA (provided by Dr G. D'Agostaro). Two identical 750-bp clones (30-2A and 30-2B) were identified in a screen of 3×10^6 recombinants and sequenced on both strands using the sequenase system (USB). They encoded the 66 amino acids at the C-terminal end of RAP30, a 3' untranslated region, and a poly(A) tail. This cDNA (30-2A) was then used to screen by hybridization⁴⁷ a human kidney cDNA library constructed in λ gt10²⁶ (provided by Dr Graeme Bell). Three identical clones with 2-kilobase (kb) inserts were isolated (9-1, 3-1, 10-1). Restriction fragments of clone 9-1 were subcloned into pEMBL18 and sequenced on both strands. The cDNA contained a 234-amino-acid open reading frame and a 3' untranslated region that was identical to the sequence of 30-2A up to the point at which poly(A) was added in 30-2A. The 3' untranslated region of the 9-1 clone continued beyond this point and did not end in poly(A) tail. The cDNA sequence that is shown has the shorter HL60-cell version of the 3' untranslated region. RAP30 that was purified by affinity chromatography and gel electrophoresis was also digested with sequencing grade trypsin or endoproteinase Asp-N (Boehringer) and the proteolytic fragments were separated by HPLC on a C8 column (RP300, Applied Biosystems). The N-terminal sequences of three tryptic fragments and one endoproteinase Asp-N fragment were determined on an Applied Biosystems Pulse Liquid Protein Sequenator (Michigan State Univ. Macromolecular Structure Facility).

```

CGGGAACGCCGCTCTTGCGCTCTCAGCGCGGCTTGCTCTTGTTCGGACGCCGCTCTCAGCGCTGGTCTGGGG
TCGCTGCTGCATCCCGACGCTCCACCGGCTGCAGACCC      ATG GCG GAG CGC GGG GAA CTC GAC TTG
MET Ala Glu Arg Gly Glu Leu Asp Leu      9

ACC GGC GCC AAA CAG AAC ACA GGA GTG TGG CTA GTC AAG GTT CCT AAA TAT TTG TCA CAG
Thr Gly Ala Lys Glu Asn Thr Gly Val Trp Leu Val Lys Val Pro Lys Tyr Leu Ser Gln      29

CAA TGG GCT AAA GCC TCT GGA AGA GGT GAA GTT GGG AAA CTG CGG ATT GCC AAG ACT CAA
Gln Trp Ala Lys Ala Ser Gly Arg Gly Glu Val Gly Lys Leu Arg Ile Ala Lys Thr Gln      49

GGA AGG ACT GAG GTG TCA TTT ACT TTG AAT GAG GAT CTT GCA AAT ATT CAT GAT ATT GGT
Gly Arg Thr Glu Val Ser Phe Thr Leu Asn Glu Asp Leu Ala Asn Ile His Asp Ile Gly      69

GGA AAA CCA GCT TCA GTC AGT GCT CCT AGA GAA CAT CCA TTT GTC TTG CAA AGT GTT GGA
Gly Lys Pro Ala Ser Val Ser Ala Pro Arg Glu His Pro Phe Val Leu Gln Ser Val Gly      89

GGA CAG ACA TTA ACA GTA TTT ACT GAG AGC TCA TCA GAT AAG CTG TCA TTG GAA GGA ATA
Gly Gln Thr Leu Thr Val Phe Thr Glu Ser Ser Ser Asp Lys Leu Ser Leu Glu Gly Ile      109

GTG GTA CAA AGA GCT GAA TGC CGA CCA GCT GCC AGT GAA AAC TAC ATG CGA TTA AAA AGA
Val Val Gln Arg Ala Glu Cys Arg Pro Ala Ala Ser Glu Asn Tyr Met Arg Leu Lys Arg      129

TTG CAA ATA GAA GAG TCT TCC AAA CCA GTG AGG CTA TCA CAA CAG CTG GAC AAA GTT GTA
Leu Gln Ile Glu Glu Ser Ser Lys Pro Val Arg Leu Ser Gln Gln Leu Asp Lys Val Val      149

ACA ACC AAT TAC AAA CCT GTT GCT AAT CAT CAA TAC AAT ATC GAA TAT GAA AGG AAA AAG
Thr Thr Asn Tyr Lys Pro Val Ala Asn His Gln Tyr Asn Ile Glu Tyr Glu Arg Lys Lys      169

AAA GAA GAC GGA AAG CGA GCT CGA GCT GAT AAA CAA CAT GTT TTA GAC ATG CTA TTT TCA
Lys Glu Asp Gly Lys Arg Ala Arg Ala Asp Lys Gln His Val Leu Asp Met Leu Phe Ser      189

GCC TTT GAG AAA CAT CAA TAC TAT AAT CTT AAG GAC TTG GTG GAC ATC ACA AAA CAA CCT
Ala Phe Glu Lys His Gln Tyr Tyr Asn Leu Lys Asp Leu Val Asp Ile Thr Lys Gln Pro      209

GTG GTG TAC CTG AAG GAA ATC TTA AAA GAA ATT GGT GTT CAG AAT GTA AAG GGA TCC ACA
Val Val Tyr Leu Lys Glu Ile Leu Lys Glu Ile Gly Val Gln Asn Val Lys Gly Ser Thr      229

AAA ACA CAT GGG AGC TGA AGCCAGAGTACAGACACTATCAAGGAGAGAAAAGAGTGACTAAGAAGACTCC
Lys Thr His Gly Ser
234

TAGCCAGCATGCTAGTGAAGGCTAGCAGCGATGCTATGCAAAAGCGCGTGAAGTCTGAAGGCTTGAACACCGTATGTT
TAATAGGGGTAAAGTGAACACTTTGATTCTCTCGGTAAATTTTAAACCTGTAAGGCTGTGAAGTTCTTAACCT
GTTTGTGAGGAGAAAGCAAGATTTATTAAGACTTAACTGTATTAACCAAGATTTTCAAGTACGAAAGAGGGG
TTGAGGATTAAGAGGGTTTCTTAAATATTAAGCTTTTAAAGTGTACAAATAGGAATTTTAAAGTGAAGACTTCT
TAGCCTTCCGCTATCAGGAGCTAGGTAATACAGATCCAAAGGAGAAATGTATAGCAAGAAAACAGCTCAACTACA
GAAACTCTTAAAGGATAAAAGCCAAAGTTCCTTATTTTGAATGTCAAGCAAGCAAGGTTTCTTTTAAAC
AGGTGATCAGTTCTTTTCTACTCAACGTTAATAAAGAGAGAGTTTGTAGTGAAAAAAGAAAAA

```

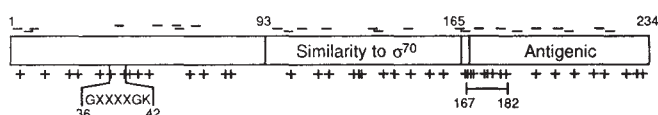
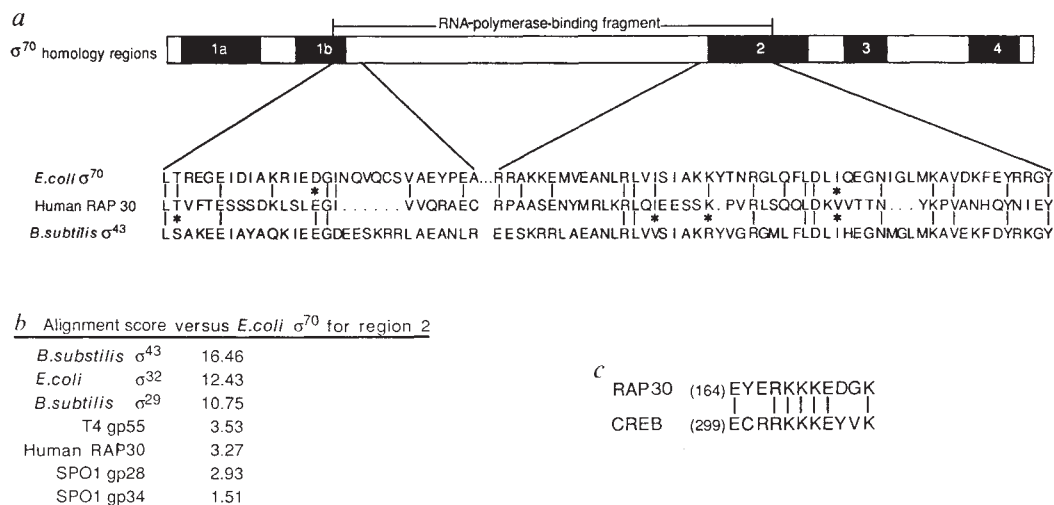


FIG. 5 Comparisons of RAP30 with σ^{70} and with the DNA-binding site of CREB. *a*, The predicted RAP30 protein sequence was compared directly with the sequence of the major sigma factor in *E. coli*, σ^{70} , using the ALIGN program in BIONET. The shaded parts of the bar represent the various regions of σ^{70} that are conserved among σ -factors³⁶. The best match between RAP30 (residues 118–165) and σ^{70} (residues 373–425) (31% identity) involved part of region 2 as indicated, whereas the second-best match (28% identity) involved part of region 1b (residues 94–117 in RAP30, residues 111–139 in σ^{70}). Regions 1b and 2 are separated by 245 amino acids in σ^{70} , but are directly adjacent in other bacterial and phage σ -factors. The line represents a tryptic fragment of σ^{70} that retains the ability to interact with the core component of RNA polymerase³⁸. *b*, Statistical analysis of the alignment that is shown for region 2 used the modified Dayhoff alignment program (ALIGN) in the Protein Identification Resource (PIR). This alignment was compared with all other possible alignments of RAP30 with 100 random permutations of the σ^{70}



sequence in this region. Other bacterial and bacteriophage σ -factors were also compared with σ^{70} in the same region, using the alignments given by Gribskov and Burgess³⁶. All scores >3 are significant. The various alignment scores are shown in rank order. Each represents the number of standard deviations by which the given alignment is greater than the average for random alignments. *c*, A possible alignment of amino acids 164–174 of RAP30 with amino acids 299–309 in the DNA-binding site of CREB³².

in the yeast *Saccharomyces cerevisiae*. Because the characteristic mammalian spacing is preserved when the yeast TATA-binding factor replaces TFIID, its mammalian counterpart, in a mammalian transcription system, it has been suggested that a different general factor determines the spacing^{5,6}. The spacing may be determined by RAP30/74 if RAP30/74 has a role in positioning RNA polymerase II, or if the precise site at which RNA polymerase II initiates transcription is determined by the precise site at which RAP30/74 melts the DNA template.

Neither initiation by eubacterial RNA polymerase nor initiation by eukaryotic nuclear RNA polymerases I and III is believed to have a requirement for ATP. We consider that the unique energy requirement for accurate initiation by RNA polymerase II¹ is at least partly explained by the ATP requirement of the DNA-helicase activity that is bound to RAP30. The assay we have used in our present studies for DNA-helicase activity is artificial in that the helicase is not positioned at a promoter by a complete initiation complex. The DNA helicase that is associated with RAP30 and initiation by a rapid-start complex

containing RNA polymerase II are inhibited by the same concentrations of Sarkosyl, and there is an exact correspondence between the nucleotides that can be used by the DNA helicase and those that can be used for accurate initiation by RNA polymerase II^{1,24}. This indicates that we have identified a DNA helicase required for accurate initiation of transcription at promoters by human RNA polymerase II.

The transition from a closed initiation complex to an open complex is a target for regulation in bacteria. In particular, the λ phage cI repressor activates transcription at the P_{rm} promoter in this way⁴⁴, and so does the phosphorylated enhancer-binding protein NR₁ of *E. coli* when it activates the *glnAp2* promoter⁴⁵. Eukaryotic transcriptional activators may, in some cases, similarly regulate the RAP30-associated DNA helicase. The availability of a cloned human RAP30 cDNA and of a simple enzymatic assay for the activity of a DNA helicase associated with RAP30 should facilitate the further analyses of protein-protein interactions and mechanistic steps involved in the initiation and regulation of transcription. □

Received 30 March; accepted 8 September 1989.

1. Bunick, D., Zandomeni, R., Ackerman, S. & Weinmann, R. *Cell* **29**, 877–886 (1982).
2. Davidson, B. L., Egly, J.-M., Mulvihill, E. R. & Chambon, P. *Nature* **301**, 630–686 (1983).
3. Parker, C. S. & Topol, J. *Cell* **36**, 357–369 (1984).
4. Nakajima, N., Horikoshi, M. & Roeder, R. G. *Molec. cell. Biol.* **8**, 4028–4040 (1988).
5. Buratowski, S., Hahn, S., Sharp, P. A. & Guarente, L. *Nature* **334**, 37–42 (1988).
6. Cavallini, B. *et al. Nature* **334**, 77–80 (1988).
7. Kadonaga, J. T., Jones, K. A. & Tjian, R. *Trends Biochem.* **11**, 20–23 (1986).
8. Matsui, T., Segall, J., Weil, P. A. & Roeder, R. G. *J. Biol. Chem.* **255**, 11992–11996 (1980).
9. Samuels, M., Fire, A. & Sharp, P. A. *J. Biol. Chem.* **257**, 14419–14427 (1982).
10. Flores, O., Maldonado, E., Burton, Z. F., Greenblatt, J. & Reinberg, D. *J. Biol. Chem.* **263**, 10812–10816 (1988).
11. Soota, M., Carthew, R. W. & Greenblatt, J. *J. Biol. Chem.* **260**, 10353–10361 (1985).
12. Reinberg, D. & Roeder, R. G. *J. Biol. Chem.* **262**, 3310–3321 (1987).
13. Zheng, X.-M., Moncollin, V., Egly, J.-M. & Chambon, P. *Cell* **50**, 361–368 (1987).
14. Fire, A., Samuels, M. & Sharp, P. A. *J. Biol. Chem.* **259**, 2509–2516 (1984).
15. Burton, Z. F., Ortalan, L. G. & Greenblatt, J. *EMBO J.* **5**, 2923–2930 (1986).
16. Burton, Z. F., Killeen, M., Soota, M., Ortalan, L. G. & Greenblatt, J. *Molec. cell. Biol.* **8**, 1602–1613 (1988).
17. VanDyke, M. W., Roeder, R. G. & Sawadogo, M. *Science* **241**, 1335–1338 (1988).
18. Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. A. *Cell* **56**, 549–561 (1989).
19. Conaway, J. W. & Conaway, R. C. *J. Biol. Chem.* **264**, 2357–2362 (1989).
20. Stahl, H., Droge, P. & Knippers, R. *EMBO J.* **5**, 1939–1944 (1986).
21. Matson, S. W. *J. Biol. Chem.* **263**, 10169–10175 (1988).
22. Goetz, S. G., Dean, F. B., Hurwitz, J. & Matson, S. J. *J. Biol. Chem.* **263**, 383–392 (1988).
23. Gelder, K. & Hoffman-Berling, H. A. *Rev. Biochem.* **50**, 233–260 (1981).
24. Sawadogo, M. & Roeder, R. G. *J. Biol. Chem.* **259**, 5321–5326 (1984).
25. Hawley, D. K. & Roeder, R. G. *J. Biol. Chem.* **260**, 8163–8172 (1985).

26. Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1194–1198 (1983).
27. Kozak, M. *Microbiol. Rev.* **47**, 1–45 (1983).
28. Moller, W. & Amons, R. *FEBS Lett.* **186**, 1–7 (1985).
29. Pabo, C. O. & Sauer, R. T. A. *Rev. Biochem.* **53**, 293–321 (1984).
30. Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609–1614 (1985).
31. Mitchell, P. J. & Tjian, R. *Science* **245**, 371–378 (1989).
32. Gonzalez, G. A. *et al. Nature* **337**, 749–752 (1989).
33. Linder, P. *et al. Nature* **337**, 121–122 (1989).
34. Ishii, S. *et al. Nucleic Acids Res.* **12**, 3333–3342 (1984).
35. Stragier, P., Parsot, C. & Bouvier, J. *FEBS Lett.* **187**, 11–15 (1985).
36. Gribskov, M. & Burgess, R. R. *Nucleic Acids Res.* **14**, 6745–6763 (1986).
37. Hellmann, J. D. & Chamberlin, M. J. A. *Rev. Biochem.* **57**, 839–872 (1988).
38. Strickland, M. S., Thompson, N. E. & Burgess, R. R. *Biochemistry* **27**, 5755–5762 (1988).
39. Hu, J. C. & Gross, C. A. *Molec. Gen. Genet.* **191**, 492–498 (1983).
40. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. *Meth. Enzym.* **91**, 524–545 (1983).
41. Chamberlin, M. J. A. *Rev. Biochem.* **43**, 721–775 (1974).
42. Travers, A. A. & Burgess, R. R. *Nature* **222**, 537–540 (1969).
43. Smale, S. T. & Baltimore, D. *Cell* **57**, 103–113 (1989).
44. Hawley, D. K. & McClure, R. W. *J. molec. Biol.* **157**, 493–525 (1982).
45. Ninfa, A. J., Reitzel, L. J. & Magasanik, B. *Cell* **50**, 1039–1046 (1987).
46. Shapiro, D. J., Sharp, P. A., Wahli, W. & Keller, M. J. *DNA*, in the press.
47. Ausubel, F. M. *et al. (eds) Current Protocols in Molecular Biology* Vol. 1 (Wiley, New York, 1987).
48. Sadowski, I., Stone, J. C. & Pawson, T. *Molec. cell. Biol.* **6**, 4396–4408 (1986).

ACKNOWLEDGEMENTS. We thank C. J. Ingles, R. Dunn, D. Fitzpatrick, L. Allison, M. Moyle and V. Watt for advice on cloning and sequencing; S. Kelly, J. Leykam and M. Corlew for help with protein sequencing; S. Matson for advice on DNA helicase assays, and C. J. Ingles, J. Segall, S. McCracken and M. Killeen for materials and for critically reading the manuscript.

Novel putative receptor tyrosine kinase encoded by the melanoma-inducing *Tu* locus in *Xiphophorus*

Joachim Wittbrodt, Dieter Adam, Barbara Malitschek, Winfried Mäueler, Friedrich Raulf, Agnes Telling, Scott M. Robertson & Manfred Schartl*

Genzentrum, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-8033 Martinsried, FRG

Malignant melanoma in *Xiphophorus* fish hybrids is caused by the activity of a dominant oncogene *Tu*. By combining genetic and molecular approaches, we have isolated the melanoma oncogene. We show that its level of expression correlates with the degree of malignancy of the tumour. The corresponding proto-oncogene is developmentally regulated. The *Tu* gene codes for a novel receptor tyrosine kinase which is closely related to the receptor for epidermal growth factor.

TUMORIGENESIS is generally considered to be controlled by oncogenes, which specify many of the malignant features of tumour cells. Oncogenes arise mainly by activation of their normal cellular counterparts, proto-oncogenes, which themselves seem to have essential roles in normal cellular physiology and are often involved in the regulation of normal cell proliferation and differentiation. By contrast, there are other genes, which are believed to be critically involved in the negative control of cell growth¹, that must be mutated at both alleles for the neoplastic phenotype of the cell to develop^{2,3}. These observations have led to the distinction between recessive oncogenes (otherwise known as tumour suppressors or anti-oncogenes) and dominant oncogenes. Although the structure and molecular function of many oncogenes are well established, the evidence that their activation is indeed the primary cause and not a consequence of tumorigenesis in the intact organism is generally circumstantial rather than direct. In *Drosophila*, genetic and molecular techniques have defined the causative role of the recessive oncogene *lethal (2) giant larvae* in tumorigenesis^{4,5}. Until now, however, comparable studies of dominant oncogenes have been limited to tumour viruses and tissue- and organ culture. Here we describe a genetic system, melanoma formation in the fish *Xiphophorus*, which allows the study of the induction and progression of malignant melanoma in vertebrates. We used this system to clone and characterize the dominant tumour-inducing gene, which we show encodes a novel putative receptor tyrosine kinase (RTK) of epidermal growth factor (EGF)-receptor type.

Genetics of melanoma formation

The platyfish (*Xiphophorus maculatus*) and the swordtail (*Xiphophorus helleri*) have a uniform grey body colouration due to small black pigment cells ('micromelanophores'). Some platyfish, however, exhibit spot patterns of melanophores that are much larger than normal ('macromelanophores')⁶. In hybrid offspring that result from the mating of spotted platyfish with non-spotted swordtails, these macromelanophore spots develop into malignant melanomas⁶⁻⁸. Genetic analyses showed that this

tumour formation depends on the abnormal regulation of a single genetic locus, the macromelanophore locus of the platyfish in hybrids⁹⁻¹². Further genetic studies led to a model for this melanoma development¹³, in which a dominant tumour gene *Tu* (equated with the macromelanophore locus) is under the control of a regulatory gene *R* which acts as a tumour suppressor. The gene *Tu* is located on sex chromosomes, whereas *R* is located on an autosome^{14,15}. In the platyfish both *Tu* and *R* are present, but they are both presumed to be absent in the swordtail. The concerted action of *Tu* and *R* results in non-proliferating macromelanophore spots. A variety of spot patterns have been identified in wild-type fish¹⁶, and the corresponding gene loci are accordingly named *Tu-Sd* (spotted dorsal pattern), *Tu-N* (nigra pattern), *Tu-Sr* (striped pattern), and so on. These loci are located either on the platyfish X or Y chromosome. In this report we mainly look at the *Tu-Sd* locus, which maps to the X chromosome. Serially repeated backcrossing of *Tu-Sd*-bearing platyfish to swordtails results in the progressive replacement of *R*-bearing platyfish chromosomes by *R*-free swordtail chromosomes (Fig. 1a). The stepwise elimination of regulatory genes is thought to lead to a corresponding stepwise deregulation of *Tu*. This leads to either benign melanoma if only one functional allele of *R* remains, or to malignant melanoma if *R* is totally absent (for review see ref. 17).

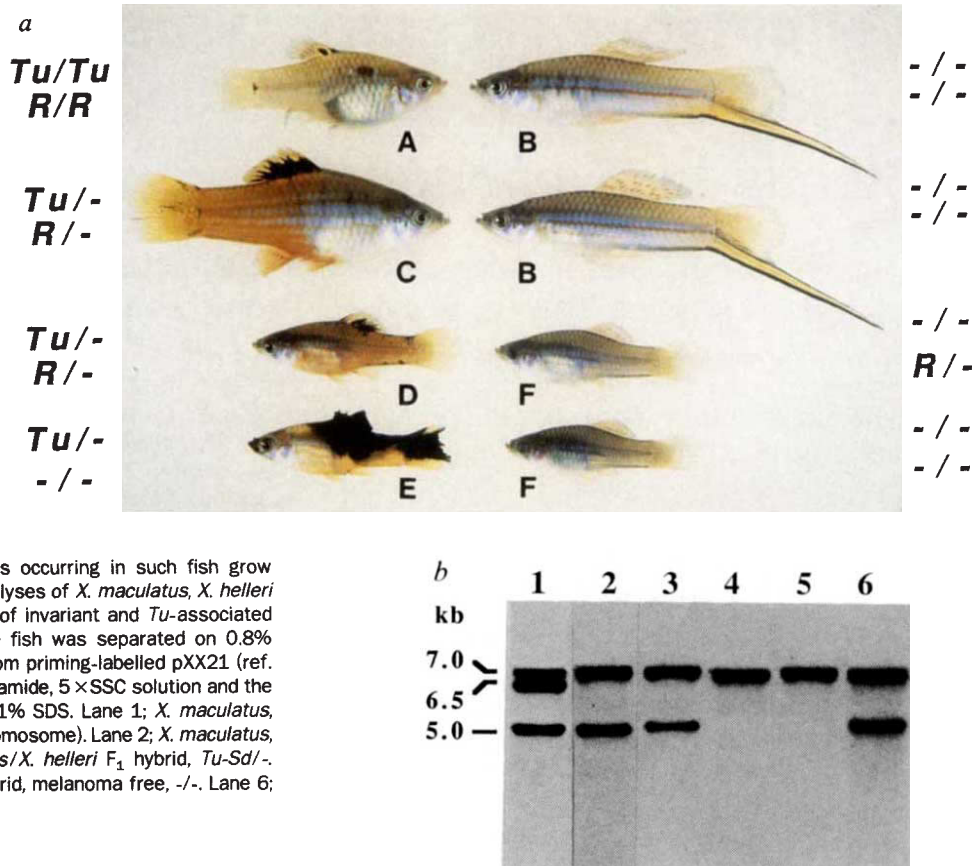
Mapping of a marker to the *Tu-Sd* gene locus

The *Tu-Sd* locus has been mapped relative to several other known loci on the X chromosome of the platyfish as follows: cen-P-1%-*Ir*-(*Pt-Dr*)-0.03%-(*Tu-Sd*)-ter, where the physical distance between the loci is represented by the recombination fraction (%) between them, as indicated (see refs 16, 18-20, and our own unpublished data). We have previously identified a restriction fragment length polymorphism (RFLP)²¹, with limited sequence similarity to a *v-erb-B* gene probe, as a diagnostic marker for sex chromosomes bearing *Tu*. We have cloned a fragment²² of the corresponding RFLP sequence of the platyfish and used it for Southern blot analyses. DNA digested with *EcoRI* gave an invariant band of 7.0 kilobases (kb), which we have called INV, in all platyfish, swordtails and their hybrids tested ($n = 135$), regardless of the presence or absence of a *Tu* locus. For fish with the X-chromosomal *Tu-Sd* locus ($n = 84$) an additional 5.0 kb band was detected, whereas for fish with the Y-chromosomal *Tu-Sr* locus ($n = 6$), a 6.5 kb band was detected in addition to the 7.0-kb band (Fig. 1b).

To further define the linkage between the *Tu-Sd* locus and the 5.0 kb RFLP marker, we analysed backcross (BC) hybrids of platyfish and swordtails for segregation; swordtails were used as the recurrent parent (BC₄-BC₇, crossed as in Fig. 1a). No recombination between the 5.0 kb RFLP marker and the *Tu-Sd* locus was observed for all 117 fish tested (Table 1). Both markers, therefore, map on the X chromosome of the platyfish with a maximal separation of 0.85 centimorgans (a recombination fraction of 1% between loci indicates a physical map distance of 1 centimorgan). Our data therefore localize the 5.0 kb marker to the *Tu-Sd*-containing region of the X chromosome. Whereas

* To whom correspondence should be addressed.

FIG. 1 a, Crossing scheme of platyfish (*X. maculatus*) and swordtails (*X. helleri*) with the corresponding *Tu/R* constellations (see text). A female *X. maculatus* (A) homozygous for the X-chromosomal macromelanophore *Tu-Sd*-gene (spotted dorsal, small spots in the dorsal fin) is mated to *X. helleri* (B), which does not carry the corresponding locus and exhibits the uniform wild-type pigmentation. The heterozygous F_1 hybrid (C) shows enhancement of the spotted dorsal phenotype. Backcrossing of the F_1 hybrid to *X. helleri* results in offspring, 50% of which have not inherited the *Tu-Sd* locus and are phenotypically like the *X. helleri* parental strain (F), and 50% of which carry the macromelanophore locus and develop melanoma. The malignancy grade ranges from very benign in some individuals (D) (phenotype like the F_1 hybrids), to highly malignant in other individuals (E). Highly malignant melanomas occurring in such fish grow invasively and are fatal. **b**, Southern-blot analyses of *X. maculatus*, *X. helleri* and their hybrids for segregation analyses of invariant and *Tu*-associated RFLP loci. *EcoRI*-digested DNA from single fish was separated on 0.8% agarose gels, blotted and probed with random priming-labelled pXX21 (ref. 22). Hybridization was at 42 °C in 50% formamide, 5 × SSC solution and the filters were washed at 68 °C in 0.1 × SSC, 1% SDS. Lane 1: *X. maculatus*, male, *Tu-Sr* (Y-chromosome)/*Tu-Sd* (X-chromosome). Lane 2: *X. maculatus*, female, *Tu-Sd*/*Tu-Sd*. Lane 3: *X. maculatus*/*X. helleri* F_1 hybrid, *Tu-Sd*/-. Lane 4: *X. helleri*, male, -/-. Lane 5: BC₄-hybrid, melanoma free, -/-. Lane 6: BC₄-hybrid with melanoma, *Tu-Sd*/-.



the invariant 7.0-kb sequence is present in all fish, the *Tu-Sd*-specific 5.0 kb sequence seems to be an additional genomic constituent of fish that have the *Tu-Sd* gene locus. By analogy we conclude that the 6.5 kb marker can be assigned to the *Tu-Sd*-containing region of the platyfish Y chromosome (Fig. 1b).

Cloning of the *Tu*-linked RFLP sequences

To clarify the relationship of the RFLP sequences to *Tu*, the invariant (INV), Y chromosomal (Y) and X chromosomal (X) *EcoRI* restriction fragments were cloned from a sub-genomic library of male platyfish (X^{Tu-Sd}/Y^{Tu-Sr}). The corresponding partial complementary DNAs, incomplete for the 5' end of the gene, were isolated from two separate cDNA libraries. The first library was generated from a melanoma that had arisen because of the presence of the X-chromosomal *Tu-Sd* locus, and the second library from a cell line (PSM; ref. 23) derived from a melanoma that had arisen because of the presence of a Y chromosomal *Tu* locus. To obtain the entire sequence of one of the cDNAs we used oligonucleotides for preparation of a primer extension library. A 2.4-kb clone, which overlapped the 1.7-kb PSM clone by 98 nucleotides, was isolated and found to contain the missing 5' sequence. The sequencing of the three

genomic *EcoRI* fragments and their comparison with the cDNA sequences revealed that each genomic clone contained the 3' portion of the same gene (Fig. 2b, c). Despite a high degree of sequence conservation throughout the exons and the introns (99%) as well as their common arrangements and sizes in homologous regions, several structural differences confirmed the independent nature of the three clones and allowed their assignment to the individual cDNAs. The three *EcoRI* fragments are therefore regarded as deriving from three different genomic loci. Although all the three loci have almost identical restriction-site maps upstream of the *SacI* site (at nucleotide 1,915) their physical maps downstream of this position differ. Most striking are the additional sequences in the loci of the INV fragment and the Y fragment of 1,349 and 1,395 bp, respectively (Fig. 2b). These additions contain intron sequences, exon y and the first five nucleotides of the large 3' exon (exon z) encoding the trailer region. These sequences are deleted from the X chromosomal locus. The cDNA clone from a melanoma that contained the X-chromosomal *Tu-Sd* locus, lacked exon y as expected, whereas the Y-chromosomal cDNA did not (Fig. 2c); this confirmed the genomic locus arrangement. The amino-acid sequence predicted from the X-chromosomal cDNA sequence indicated that the deletion does not change the reading frame, and thus represented a polypeptide with an internal deletion of 35 amino acids (see below). Neither cDNA extended to the poly(A) tail, although all three genomic loci contain a polyadenylation-site signal at corresponding positions. This would add another 229 nucleotides downstream of the cDNA sequence (Fig. 3).

RFLPs are a structural component of the *Tu* locus

That the sex-chromosomal genes are actually encoded by the *Tu* locus is shown by the presence of 'loss of function' *Tu* alleles. We analysed two spontaneous *Tu* mutants that do not exhibit a single macromelanophore and have lost the capacity to develop malignant melanoma in backcross hybrids. In one mutant

TABLE 1 Segregation of *Tu-Sd* and the 5.0 kb marker sequence

	<i>Tu-Sd</i>	5.0-kb RFLP marker	n
Parental class 1	<i>Tu-Sd</i>	5.0-kb RFLP marker	76
Parental class 2	+	+	41
Recombinant class 1	<i>Tu-Sd</i>	+	0
Recombinant class 2	+	5.0-kb RFLP marker	0
Total = 117			

* The 5.0-kb *EcoRI* fragment is additional to the invariant 7.0-kb fragment of wild-type fish, therefore lack of the 5.0-kb RFLP marker is regarded as the wild-type situation (+).

derived from the Y-chromosomal *Tu-Sr* locus, the 6.5-kb fragment is totally-deleted (Fig. 2d). In another mutant, a rearranged *EcoRI* fragment of 12 kb instead of the expected 5.0 kb, in addition to the invariant 7-kb fragment, was detected with the RFLP probe (Fig. 2d). Because we have so far detected no intra-strain polymorphism for the X- or Y-chromosomal marker sequences (see Table 1, ref. 21, and unpublished data) this 12-kb fragment was cloned and analysed. Restriction mapping and sequence analysis precisely defined an insertion of ~7 kb between base-pairs 1,652 and 1,653 in exon w (Fig. 2e). This insertion, therefore, interrupts the coding sequence of the

X-chromosomal gene. The result of this alteration is the loss of the *Tu*-phenotype. We conclude therefore that the X-chromosomal 5.0-kb marker sequence is the 3' part of the *Tu-Sd* gene locus, the Y-chromosomal 6.5-kb marker by analogy is the homologous part of *Tu-Sr*, whereas the 7.0-kb marker represents a gene found in all fish independent of sex-chromosomal *Tu* loci.

A novel putative receptor tyrosine kinase

The almost full-length cDNA clone isolated from the PSM cell line contains a single long open-reading frame (ORF) of 3498 nucleotides, which starts with ATG and terminates with an

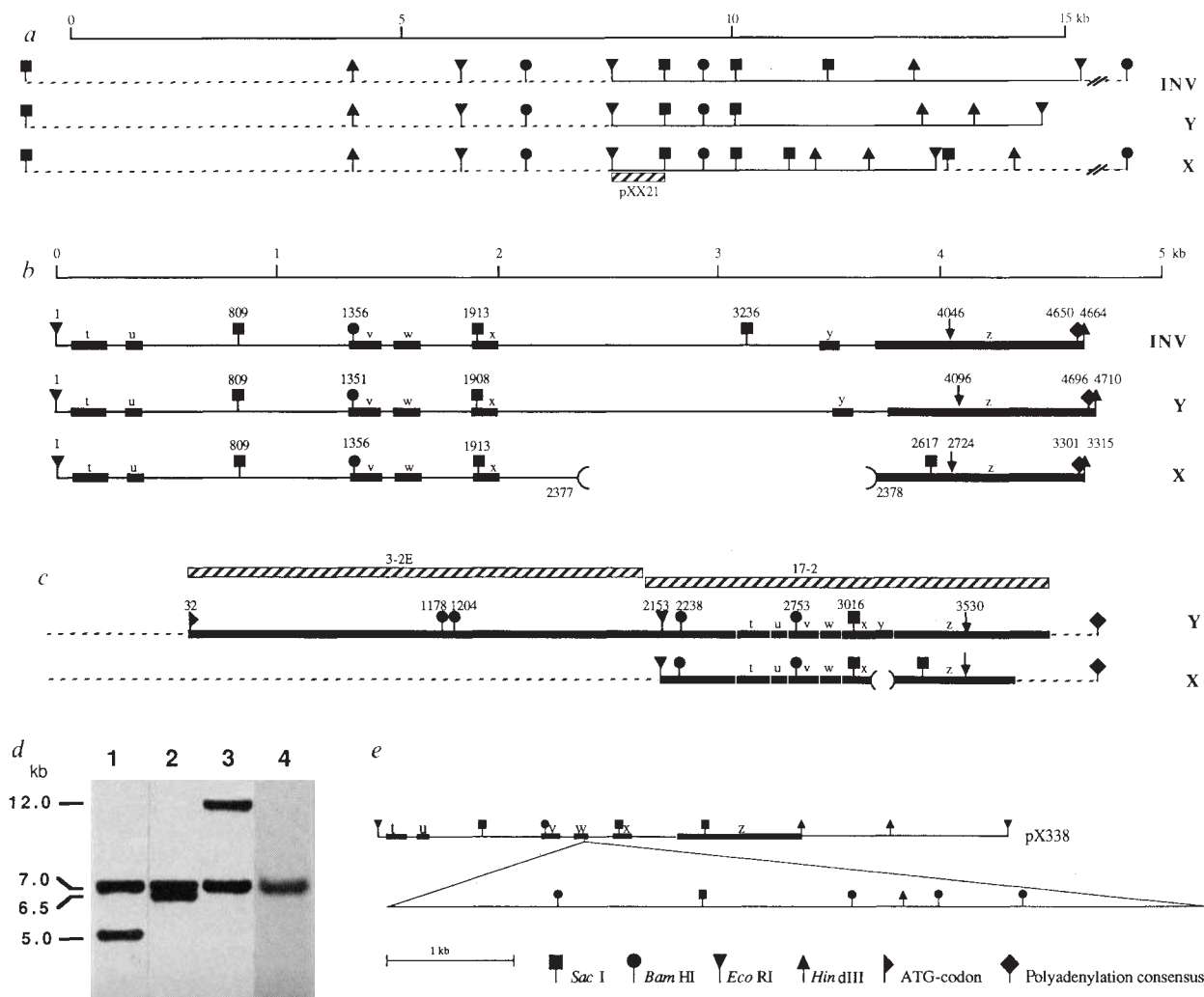


FIG. 2 Partial structure of the genomic region of INV-, Y- and X-linked RFLP loci and corresponding cDNAs. *a*, Restriction map of the genomic loci. *b*, Exon-intron arrangements in the 3' region of INV-, Y- and X-genomic loci. Differing positions of exons, stop codons, restriction sites and polyadenylation consensus sites downstream of nucleotide 1,277 between INV-, X- and Y loci result from sequence differences in the corresponding trailer and intron regions. *c*, Comparison of X- and Y-loci transcripts. Sequenced cDNAs are marked in black. *d*, RFLP analysis of the loss-of-function mutants. Southern-blot analysis of *EcoRI*-digested DNA of *X. maculatus*/*X. helleri* BC hybrids (recurrent parent *X. helleri*) carrying the parental wild-type *Tu-Sd*-bearing X chromosome (BC₁, lane 1), *Tu-Sr*-bearing Y chromosome (BC₂, lane 2) and the two independent mutant chromosomes (BC₂, lanes 3 and 4). For methods see Fig. 1. *e*, Restriction map of the rearranged 12-kb fragment from the loss-of-function mutant (see Fig. 2d, lane 3). Relevant stop codons are marked by arrows; deleted fragments are bounded by semi-circles; cloned fragments are indicated by solid lines; structural data obtained by Southern- or Northern-blot hybridizations are shown by broken lines; probes used for screening shown by hatched boxes.

METHODS. INV-, Y- and X-linked fragments were cloned as follows: High molecular weight genomic DNA from male *X. maculatus* (stock Rio Jamapa

X. Tu-Sd, *Y Tu-Sr*) was prepared by standard methods⁴¹. The DNA was digested to completion with *EcoRI*. Fragments were separated on a LMP-agarose gel. Gel fractions with sizes of ~5 kb, 6.5 kb and 7 kb were cut out; purified DNA was ligated into λ gt10 arms and packaged. The sub-libraries were screened with a *v-erb B* probe (600-bp *Bam*HI fragment D of pAE11 ref. 42) as well as with pXX21 (ref. 22). Hybridization conditions were 42 °C, 40% formamide, 5 × SSC (wash at 60 °C, 2 × SSC) for pAE11, and as described in Fig. 1 for pXX21. Genomic *EcoRI*-*Hind*III sub-fragments of INV, Y and X and the corresponding cDNAs were subcloned into pBluescript (Stratagene). For cloning of the 12-kb fragment from the loss-of-function mutant, DNA from a BC₂ hybrid containing the mutant chromosome was cloned as outlined above into λ EMBL4. The subgenomic library was screened with the 17-2 the probe (Fig. 2c) under high stringency conditions. Sequencing: nested deletions were generated using either Exonuclease III (ref. 43), or T4 DNA polymerase⁴⁴ (using oligonucleotides adapted for pBluescript). Sequences were determined by dideoxy chain-termination sequencing using Sequenase (US Biochemical Corporation) or T7 DNA polymerase (Pharmacia). Sequence analysis was performed using the GCG sequence analysis software package⁴⁵.

NATURE • VOL 341 • 5 OCTOBER 1989

XER protein has in common with the HER- and the CER proteins that are not common to the *Xmrk*-encoded protein (Fig. 4a). Within this region, *Xmrk* has only 76% identity to the gene that encodes XER, and the two proteins that these genes encode have 83% identity. Sequence comparisons shows that the chicken and human EGF-receptor proteins have extended motifs in common throughout their extracellular domains and C-terminal tails, which are divergent from the corresponding sequences predicted for the *Xmrk*-encoded protein (Fig. 4a). The *Xmrk*-encoded protein is not therefore, the fish EGF-receptor homologue, but a novel putative receptor tyrosine kinase.

Elevated expression in melanoma

To investigate the physiological function of *Xmrk* we studied its expression pattern during embryogenesis, in normal adult tissue and in melanomas, induced by sex-chromosomal *Tu* loci, that differ in their degree of malignancy. Northern-blot hybridizations of RNA obtained from swordtail embryos containing only the INV *Xmrk* locus revealed a single *Xmrk* transcript of 5.8 kb (Fig. 5a). This transcript was highly abundant as a maternal messenger RNA in unfertilized eggs (Fig. 5a, lane 1). The level decreased during early development (cleavage to neurula stages), increased during organogenesis and from then on decreased. In 4-week-old neonate fish, only barely detectable hybridization signals were obtained (Fig. 5a, lane 10). In normal organs of adult fish the transcript was seen at low levels in gills, fins and skin but not in any other organ tested (data not shown).

In all melanomas that originated from six different *Tu* alleles from hybrid fish and in the melanoma cell line (PSM)²³, high levels of *Xmrk* transcripts were found (Fig. 5b). The RNA preparations included both early- and late-onset melanomas²³, as well as amelanotic tumours from fish homozygous for the albino (*a*) gene. Two large transcripts were seen, one that corresponds in size to the transcript from the INV locus (5.8 kb; see above) and a smaller transcript (4.7 kb) which is specific for the melanoma (Fig. 5b). Shorter transcripts (less than 3 kb) represented either specific degradation products, or transcripts of different length, but from the same gene, generated by differential splicing, alternative transcription start sites or alternative polyadenylation signals. The steady-state level of the 4.7 kb mRNA and the shorter transcripts was highly correlated to the degree of malignancy of the melanoma. This was most obvious for tumours due to the *Tu*-*Sd* allele, for which back-cross genotypes of defined malignancy were available. In extremely benign lesions, which occur in the presence of one intact allele of the melanoma suppressor gene *R*, only very low amounts of *Xmrk* mRNA were seen (Fig. 5c, lane 9) compared with the amounts seen in more malignant melanoma from fish that lacked both alleles of *R* (Fig. 5c, lane 10). In addition, fish that were heterozygous for *a*, which further enhances melanoma malignancy, had even higher transcript levels (Fig. 5b, lane 11). Expression of the INV locus was at a similarly low level in all tumours irrespective of their state of malignancy. The observation of a 5.8 kb INV *Xmrk* transcript in several normal tissues and the absence of the tumour-specific 4.7 kb transcripts indicates that the INV locus bears the *Xmrk* proto-oncogene.

Discussion

We have made use of the *Xiphophorus* melanoma system to study malignant tumours by both genetic and molecular techniques. We have cloned the oncogene *Tu* of *Xiphophorus*, known to induce malignant melanoma in the absence of the anti-oncogene *R*, and found that it encodes a novel transmembrane tyrosine kinase. We have called this gene *Xmrk*. In addition to the oncogenic *Xmrk* loci on the X- and Y chromosomes of platyfish, we found that another copy of *Xmrk* at the INV locus is a normal constituent of the genome of all *Xiphophorus* fish, independent of the melanoma oncogene. This copy is probably the *Xmrk* proto-oncogene, which could have a normal physio-

logical function during development and in specific tissues.

The proposed structure of the protein encoded by *Xmrk* and its similarity to the subclass I growth-factor-receptor tyrosine kinases indicates that it is a transmembrane protein with an extracellular domain acting as a receptor, and a cytoplasmic tyrosine-kinase domain. RTKs are normally activated by the binding of a specific ligand which results in the phosphorylation of specific cellular proteins, thus producing a pleiotropic response, including proliferation and/or differentiation²⁸. It is probable that *Xmrk* participates in such a signal-transduction mechanism. Although the predicted protein encoded by *Xmrk* shows structural similarity to the EGF-receptor of higher vertebrates, our data show that it does not represent the fish EGF-

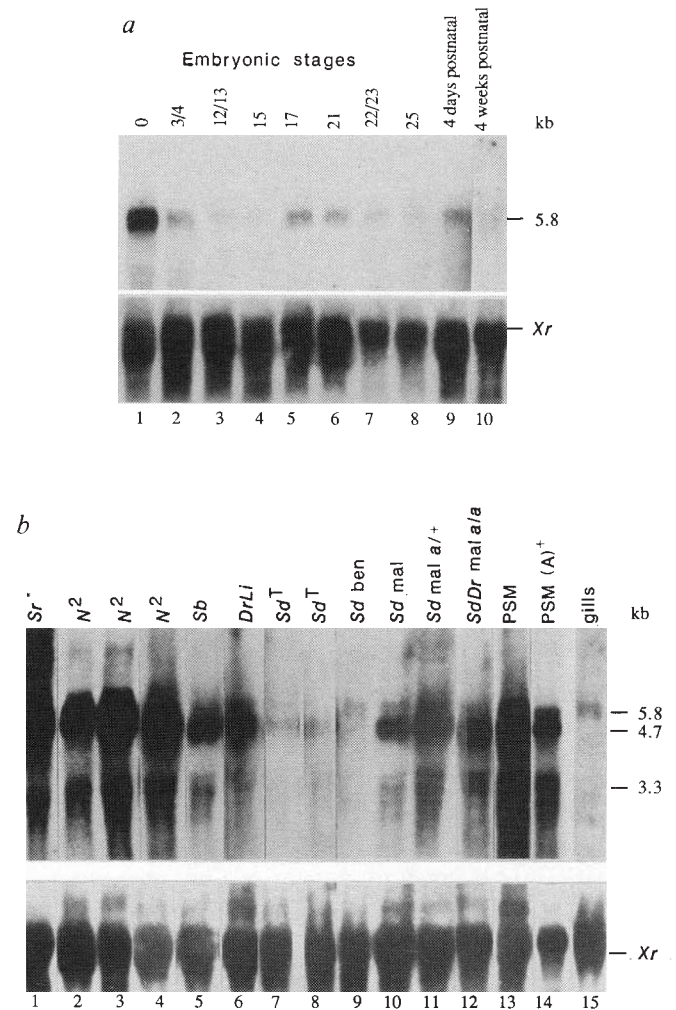


FIG. 5 Northern-blot analyses. *a*, Profile of *Xmrk* expression during embryonic development (for staging see ref. 47) (lanes 1–8), and at two different postnatal times (lanes 9 and 10) in fish containing only *Xmrk* at the INV locus. In each case, 20 μ g of total RNA was analysed by Northern-blot hybridization. *b*, *Xmrk* expression in melanomas and in gills (lane 15). Total RNA (20 μ g) from single tumours (lanes 1–4, 6, 11) or pooled melanoma biopsies (lanes 5, 7–10, 12) and a melanoma cell line (lanes 13 and 14) were analysed. PSM, melanoma cell line; *a*, albino gene; ben, benign; mal, malignant; all other abbreviations refer to melanomas due to different *Tu* alleles (lanes 1–5, Y-chromosomal; lanes 6, 9–12, X-chromosomal; lanes 7, 8, autosomal); *Sr*, mutant striped; *N*², nigra-2; *Sb*, spotted belly; *DrLi*, mutant lineatus; *Sd*^T, spotted dorsal translocation; *Sd*, spotted dorsal. METHODS. RNA preparation, Northern blots and hybridizations for homologous probes were performed as described³⁸. Probes used were random-priming-labelled 17–2 DNA (*b*) or a riboprobe of 17–2 (*a*). For size calibration a RNA ladder (BRL, Bethesda) was included. As a control for the amount of RNA blotted onto the membranes, the ubiquitously expressed *Xr* gene⁴⁸ was used as a hybridization probe.

receptor. The *Xmrk*-protein therefore, probably does not bind the fish EGF homologue. The *HER2/neu* gene product, which so far in mammals is the most closely related protein to the EGF-receptor, does not recognize EGF as its physiological ligand but binds to an as yet unidentified molecule³¹⁻³³. Further experiments are needed to identify the possible ligand of the *Xmrk*-gene product RTK.

No satisfactory candidate for a RTK substrate has been identified³⁰. The normal physiological function of the *Xmrk* encoded RTK is therefore beyond speculation, although the abundant *Xmrk* maternal message is intriguing with regard to recent reports that RTKs, like the *Drosophila* EGF-receptor and the *torso*-gene product, function as maternal-effect genes during early embryonic development³⁴⁻³⁶.

The existence of the sex-chromosomal oncogenic *Xmrk* loci as additional genomic constituents in the platyfish poses the question of their evolutionary origin. It is conceivable that an allelic or duplicated copy of the INV-locus *Xmrk* gene was translocated to the sex-chromosomal pigmentary loci thereby acquiring an oncogenic potential. This is supported by the identification of macromelanophore pigment patterns in several poeciliid fish species and even in the genus *Xiphophorus* that never give rise to tumour formation after the hybridization (ref. 10, and our own unpublished data). In purebred platyfish the malignant phenotype is suppressed because of the presence of the *R* gene. After crossing-conditioned elimination of the *R* gene, melanoma develops. We have shown that the steady-state levels of sex-chromosomal *Xmrk* transcripts correlate with the genetically determined malignancy of the tumour, raising the possibility that transformation is affected by an increase in the level of *Xmrk* transcription. Such a situation is reminiscent of *c-myc* activation in Burkitt's lymphoma (reviewed in ref. 37) in which inappropriate regulation of *myc* is a consequence of a translocation.

Over-expression of RTKs of subclass I in tissue-culture leads potentially to oncogenesis, and increased levels are frequently, associated with certain human malignancies (for a review see

ref. 30). It should be noted, however, that in these tumours increased expression is mostly due to gene amplification, whereas in *Xiphophorus* melanomas no evidence for an amplification of *Xmrk* or related genes was obtained^{21,38}. Recent studies indicate that in addition to over-expression of RTKs (in conjunction with autocrine ligand production), structural alterations may have a key role in the initiation as well as the progression of certain neoplasms (for a review see ref. 30). The *Xmrk* melanoma-specific transcripts are ~1 kb shorter than the transcript from the proto-oncogenic *Xmrk* at the INV locus. This points to such a structural alteration, possibly associated with the proposed translocation event. The full significance of the differing transcripts awaits further elucidation. The C-terminal deletion observed for the *Xmrk*-gene products encoded by the X-chromosomal locus seems not to affect its transforming capability in general, because in this region the potentially oncogenic *Xmrk* at the Y-chromosomal locus contains an identical exon as the proto-oncogenic *Xmrk* at the INV locus. This difference, however, may account for an observed variability in the malignancy of the resulting melanoma; the X-chromosomal *Tu-Sd* locus is responsible for much more severe tumorous lesions than the Y-chromosomal *Tu-Sr* locus. For the subclass I RTKs of higher vertebrates, a variety of minor or severe alterations of the C-terminal domains do not alter transforming potentials in general, but confer higher sensitivity to ligand-induced mitogenesis than that found in the wild-type situation^{39,40} (for a discussion see ref. 30).

We propose that the melanoma oncogene of *Xiphophorus* encodes an activated version of a novel putative RTK, whose oncogenic potential is suppressed in wild-type platyfish. The suitability of fish for developmental biological studies and the availability of numerous mutants and wild-type alleles that affect the melanoma oncogene itself (or the phenotype of the resulting melanoma) makes *Xiphophorus* a unique *in vivo* system for studies of RTK-mediated oncogenesis as well as possible mechanisms for the suppression of this process. □

Received 29 June; accepted 5 September 1989.

1. Friend, S. H., Dryja, T. P. & Weinberg, R. *New Engl. J. Med.* **318**, 618-622 (1988).
2. Ponder, B. *Nature* **335**, 400-402 (1988).
3. Knudson, A. G. *Cancer Res.* **45**, 1437-1443 (1985).
4. Mechler, B. M., McGinnis, W. & Gehring, W. J. *EMBO J.* **4**, 1551-1557 (1985).
5. Jacob, L. et al. *Cell* **50**, 215-225 (1987).
6. Gordon, M. *Genetics* **12**, 253-283 (1927).
7. Häussler, G. *Klin. Wschr.* **7**, 1561-1562 (1928).
8. Kosswig, C. Z. *Indukt. Abstammungs-Vererbungsl.* **47**, 150-158 (1928).
9. Gordon, M. *Ann. N.Y. Acad. Sci.* **71**, 1213-1222 (1958).
10. Atz, J. W. *Zoologica* **47**, 153-181 (1962).
11. Kosswig, C. *Zool. Anz.* **175**, 21-50 (1965).
12. Zander, C. D. *Mitt. Hamburg. Zool. Inst.* **66**, 241-271 (1969).
13. Ahuja, M. R. & Anders, F. *Prog. Exp. Tumour Res.* **20**, 380-397 (1976).
14. Morizot, D. C. & Siciliano, M. J. *J. Natn. Cancer Inst.* **71**, 809-813 (1983).
15. Ahuja, M. R., Schwab, M. & Anders, F. *J. Hered.* **71**, 403-407 (1980).
16. Kallmann, K. D. in *Handbook of Genetics* (ed. King, I.) 81-132 (Plenum, New York, 1975).
17. Anders, F., Scharf, M., Barnekow, A. & Anders, A. *Adv. Cancer Res.* **42**, 191-275 (1984).
18. Kallmann, K. D. *Zoologica* **55**, 1-16 (1970).
19. Gordon, M. *Genetics* **22**, 376-392 (1937).
20. Kallmann, K. D. & Atz, J. W. *Zoologica* **51**, 107-135 (1966).
21. Scharf, M. *Genetics* **119**, 679-685 (1988).
22. Adam, D., Wittbrodt, J., Telling, A. & Scharf, M. *Nucleic Acids Res.* **16**, 7212 (1988).
23. Wakamatsu, Y. *Dev. Growth Differ.* **22**, 731-740 (1980).
24. Heijne, G. v. *J. Molec. Biol.* **184**, 99-105 (1985).
25. Heijne, G. v. *Nucleic Acids Res.* **14**, 4683-4690 (1986).
26. Klein, P., Kanehisa, M. & De Lisi, C. *Biochem. Biophys. Acta* **815**, 468-476 (1985).
27. Ullrich, A. et al. *Nature* **309**, 418-425 (1984).
28. Hunter, T. & Cooper, J. A. *A. Rev. Biochem.* **54**, 897-930 (1985).

29. Hanks, S. K., Quinn, A. M., & Hunter, T. *Science* **241**, 42-52 (1988).
30. Yarden, Y. & Ullrich, A. *A. Rev. Biochem.* **57**, 443-478 (1988).
31. Akiyama, T. et al. *Molec. Cell Biol.* **8**, 1019-1026 (1988).
32. King, C. R. et al. *EMBO J.* **7**, 1647-1651 (1988).
33. Stern, D. F. & Kamps, M. P. *EMBO J.* **7**, 995-1001 (1988).
34. Sprenger, F., Stevens, L. M. & Nüsslein-Vollhard, C. *Nature* **338**, 478-483 (1989).
35. Price, J. V., Clifford, R. J. & Schupbach, T. *Cell* **56**, 1085-1092 (1989).
36. Schejter, E. D. & Shilo, B.-Z. *Cell* **56**, 1093-1104 (1989).
37. Bornkamm, G. W., Polack, A. & Eick, D. in *Cellular Oncogene Activation* (ed. Klein, G.) 223-273 (Marcel Dekker, New York, 1988).
38. Mäuer, W., Raulf, F. & Scharf, M. *Oncogene* **2**, 421-430 (1988).
39. Honegger, A. et al. *EMBO J.* **7**, 3045-3052 (1988).
40. Khazaei, K. et al. *EMBO J.* **7**, 3061-3071 (1988).
41. Blin, N. & Stafford, W. *Nucleic Acids Res.* **3**, 2303-2308 (1976).
42. Vennström, B., Fanshiv, L., Moscovici, C. & Bishop, J. M. *J. Virol.* **36**, 575-585 (1980).
43. Henikoff, S. *Gene* **28**, 351-359 (1984).
44. Dale, R. M. K., McClure, B. A. & Houchins, J. P. *Plasmid* **13**, 31-40 (1985).
45. Devereux, J. R. et al. *Nucleic Acids Res.* **12**, 387-395 (1984).
46. Lax, I. et al. *Molec. Cell Biol.* **8**, 1970-1978 (1988).
47. Tavalga, W. N. *Bull. Am. Mus. Nat. Hist.* **94**, 161-230 (1949).
48. Schwab, M. *Molec. Gen. Genet.* **188**, 410-417 (1982).

ACKNOWLEDGEMENTS. We thank H. Jäckle for discussions and critical comments on the manuscript. This work was supported by grants to M.S. from the Bundesministerium für Forschung und Technologie through 'Schwerpunkt: Grundlagen und Anwendungen der Gentechnologie', the Deutsche Forschungsgemeinschaft through 'Schwerpunkt: Klassische und molekulare Tumorzitogenetik' and the Max-Planck-Gesellschaft, D.A. is a fellow of the Fonds der Chemischen Industrie.

Anisotropic ionizing radiation in NGC5252

Clive Tadhunter* & Zlatan Tsvetanov†

* Royal Greenwich Observatory, Madingley Road, Cambridge CB3 0EZ, UK

† European Southern Observatory, Karl-Schwarzschild-Strasse-2, D-8046, Garching bei München, FRG

THE different kinds of active galactic nuclei (AGNs) and quasars may be fundamentally similar, and owe their observed differences to anisotropy in the distribution of their emitted radiation: apparently different objects may then be the same thing seen from different viewpoints. In particular, it has been proposed that narrow-line Seyfert 2 galaxies are Seyfert 1 galaxies for which the broad-line and continuum emitting regions are hidden by obscuring material along our line of sight^{1,2}. Here we present narrow-band imaging observations of the Seyfert 2 galaxy NGC5252 which strongly support such unifying hypotheses. Images in the light of the [O III] line at $\lambda = 5,007 \text{ \AA}$ reveal a sharply defined bi-conical structure extending to a maximum radius of $\sim 33 \text{ kpc}$ ($H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$) along the radio-emission axis. This is just the type of structure expected if the ionizing-radiation field from the nucleus is anisotropic. A map of the ratio of the [O III] and H α lines shows a central band perpendicular to the cones that we identify as the disk or torus responsible for collimating the radiation field.

Some of the best evidence that the ionizing continuum is radiated anisotropically by active galactic nuclei is provided by observations of the surrounding extended emission-line regions (EELR). Long-slit spectroscopic and narrow-band imaging observations show that, where the emission lines are extended, they are often more extended along or near the radio-emission axis than in the perpendicular direction^{3,4}. Moreover, in many objects the highest ionization EELR are found along the radio-emission axis, and for these one commonly finds that a power-law extrapolation of the observed nuclear continuum is incapable of providing enough ionizing photons to produce the observed emission-line luminosities⁴⁻⁶. These results apply across the whole range of radio powers, from the relatively low radio-power Seyfert galaxies to the most powerful radio galaxies and quasars.

Although the link between the collimation direction of the radiation field and the radio axis is clear, the collimation mechanism remains uncertain. Several are possible, including radiation of the continuum emission by an accretion disk, leading to a $\cos \theta$ dependence of the continuum emission on the angle θ to the axis of the disk⁷, relativistic beaming^{8,9}, and simple obscuration of the radiation field by a disk or torus^{1,2} of gas and dust.

To examine the anisotropy in more detail we have obtained images and spectra of NGC5252, a Seyfert 2 galaxy which was known from previous observations to have bright emission lines preferentially extended along the radio-emission axis³. Some basic properties of NGC5252 are listed in Table 1. The observations were taken using the European Southern Observatory Faint-Object Spectrograph and Camera (EFOSC) on the ESO 3.6-m telescope at La Silla, Chile in March 1989. We report the spectroscopic observations in detail elsewhere (C.T. and Z.T. manuscript in preparation). Here we concentrate on the images that were taken with interference filters ($\Delta\lambda \approx 60 \text{ \AA}$) centred on redshifted positions of [O III], H α and two line-free continuum bands ($\lambda_c = 5,210 \text{ \AA}$ and $\lambda_c = 7,027 \text{ \AA}$).

A striking aspect of the continuum-subtracted [O III] image (Fig. 1a) is the elongated structure extending in a series of shells out to a maximum radius of $\sim 50 \text{ arcsec}$ ($\sim 33 \text{ kpc}$) to the north and $\sim 45 \text{ arcsec}$ ($\sim 30 \text{ kpc}$) to the south of the nucleus. This structure is consistent with a bi-conical geometry, with the long axis of the cones roughly aligned with the radio-emission axis. Our long-slit spectra show the EELR within the structure have

TABLE 1 Basic properties of NGC5252

<i>z</i>	0.0234	
Morphology	S0	ref. 21
M_B (15 arcsec aperture)	-20.4	ref. 15
P_{6cm} ($\text{W Hz}^{-1} \text{ sr}^{-1}$)	3.4×10^{21}	ref. 22
$L_{[O III]}$ (erg s^{-1})	1.4×10^{42}	
PA of major axis of the host S0	15°	
PA of [O III]-cone major axis	167°	
PA of central band	75°	
PA of radio-emission axis	170°	ref. 3

a high ionization state consistent with photoionization by a continuum with a substantial extreme ultraviolet (EUV) component, but inconsistent with photoionization by normal OB stars¹⁰. The opening angle of the cones is $\sim 75^\circ$. It is just the type of structure expected if the interstellar medium (ISM) in the host galaxy is illuminated by an anisotropic radiation field from the nucleus. As such, it perhaps represents the best example of radiation cones yet seen in Seyfert galaxies.

In contrast, the continuum image (Fig. 1b) reveals a bulge-dominated structure typical of S0 galaxies. The relationship between this and the [O III] cones is unclear because the major axis of the S0 disk (PA $\approx 15^\circ$) is misaligned from the cone axis by $\sim 30^\circ$. To search for evidence of an ISM associated with the S0 disk we have divided our two continuum images to form a colour map. This shows redder colours in a linear structure parallel to the S0 major axis which are consistent with obscuration by dust in the galaxy disk and therefore provide preliminary evidence that NGC5252 contains an ISM that is not photoionized by the intense component of the nuclear radiation field. The origin of the [O III]-emitting gas—much of which lies outside the galaxy disk and has a shell-like morphology—remains uncertain. Perhaps the most likely explanation is that the shells represent the illuminated sectors of ring structures that surround most of the visible galaxy. This would make NGC5252 similar to several other S0 galaxies for which large-radius ring structures have been observed in H I observations¹¹.

A clue to the nature of the radiation collimation mechanism is provided by the sharpness of the edges of the emission-line cones (Fig. 1a)—such sharp definition would not be expected if the continuum had the tangential dependence predicted for emission from an accretion disk or for relativistic beaming, but it is consistent with obscuration by a disk. Indeed, some evidence for such an obscuring region is apparent in the [O III]/H α line-ratio map presented in Fig. 1c. This shows a band of reduced [O III]/H α roughly perpendicular to the emission-line cones, which can either be interpreted as a region of higher obscuration or as a region of lower ionization. In either case, the alignment of the band leads us to associate it with the central disk or torus responsible for collimating the radiation field. We note that evidence for a similar band has recently been found for the nearby Seyfert 1 galaxy NGC4151¹². Interestingly, the major axis of the band is misaligned from the major axis of the host S0 galaxy by $\sim 60^\circ$. The position angles of the various components in NGC5252 are compared in Table 1.

Apart from the emission-line structure, we can use the line fluxes derived from our images to test the proposed unification of Seyfert 1 with Seyfert 2 galaxies^{1,2}. From simple recombination arguments, the total H α luminosity ($L_{H\alpha}$) can be related to the monochromatic continuum luminosity that an observer in the cone would infer assuming an isotropic continuum (L_λ). For a power-law optical-to-X-ray continuum ($F_\nu \propto \nu^\alpha$) we have:

$$L_\lambda = \left(\frac{-\alpha}{55f(1 - \cos(\theta/2))} \right) \left(\frac{\lambda}{\lambda_0} \right)^{-(2+\alpha)} L_{H\alpha} \quad (1)$$

where f is the fraction of ionizing photons in the cones that are absorbed by the gas, λ_0 is the wavelength of the Lyman limit,

θ is the opening angle of the cones and Case B recombination for an electron temperature of 15,000 K has been assumed. Our images give $L_{\text{H}\alpha} > 3.3 \times 10^{41} \text{ erg s}^{-1}$, after making a 50% correction for [N II] emission in the filter bandpass. Substituting this into the equation with $f=1$ and the appropriate value of θ we obtain $L_{4,400 \text{ \AA}} > 6 \times 10^{39} \text{ erg s}^{-1} \text{ \AA}^{-1}$ (or $M_B < -19.5$) for $\alpha = -1.0$, and $L_{4,400 \text{ \AA}} > 6 \times 10^{40} \text{ erg s}^{-1} \text{ \AA}^{-1}$ (or $M_B < -22.0$) for $\alpha = -2.0$ (this range of α is compatible with the ionizing continuum inferred from the emission-line ratios¹⁰). Although these luminosities are likely to be substantially underestimated because f is probably much smaller than unity, they are comparable with the nuclear luminosities of Seyfert 1 galaxies, but are somewhat greater than the nuclear luminosities of Seyfert 2 galaxies^{13,14}. Further, the aperture photometry¹⁵ and the lack of dilution of the stellar absorption features near the nucleus in our long-slit spectra together suggest that the nuclear non-stellar luminosity of NGC5252 itself (inferred from our line of sight)

is smaller than these estimates. Hence, our observations are consistent with the hypothesis that NGC5252 would look much brighter—possibly like a Seyfert 1 galaxy—to an observer within the cone.

NGC5252 has a relatively low total radio power, so it is also of interest to compare its properties with those of more powerful radio sources ($P_{6\text{cm}}^{\text{tot}} > 10^{22} \text{ W Hz}^{-1} \text{ sr}^{-1}$). We note that the extent of the emission lines, the high ionization spectrum and the emission-line luminosity in NGC5252 are typical of the EELR around powerful radio galaxies and quasars (refs 16–18 and S. de Serego Alighieri, manuscript in preparation). It is more difficult to compare the opening angle θ , because few powerful radio galaxies have such well defined cones. The statistics of the alignment between the very extended emission-line regions and the radio-emission axes in powerful radio sources¹⁹, however, are consistent with the value of θ we have derived here. Similar values of θ are also required by schemes

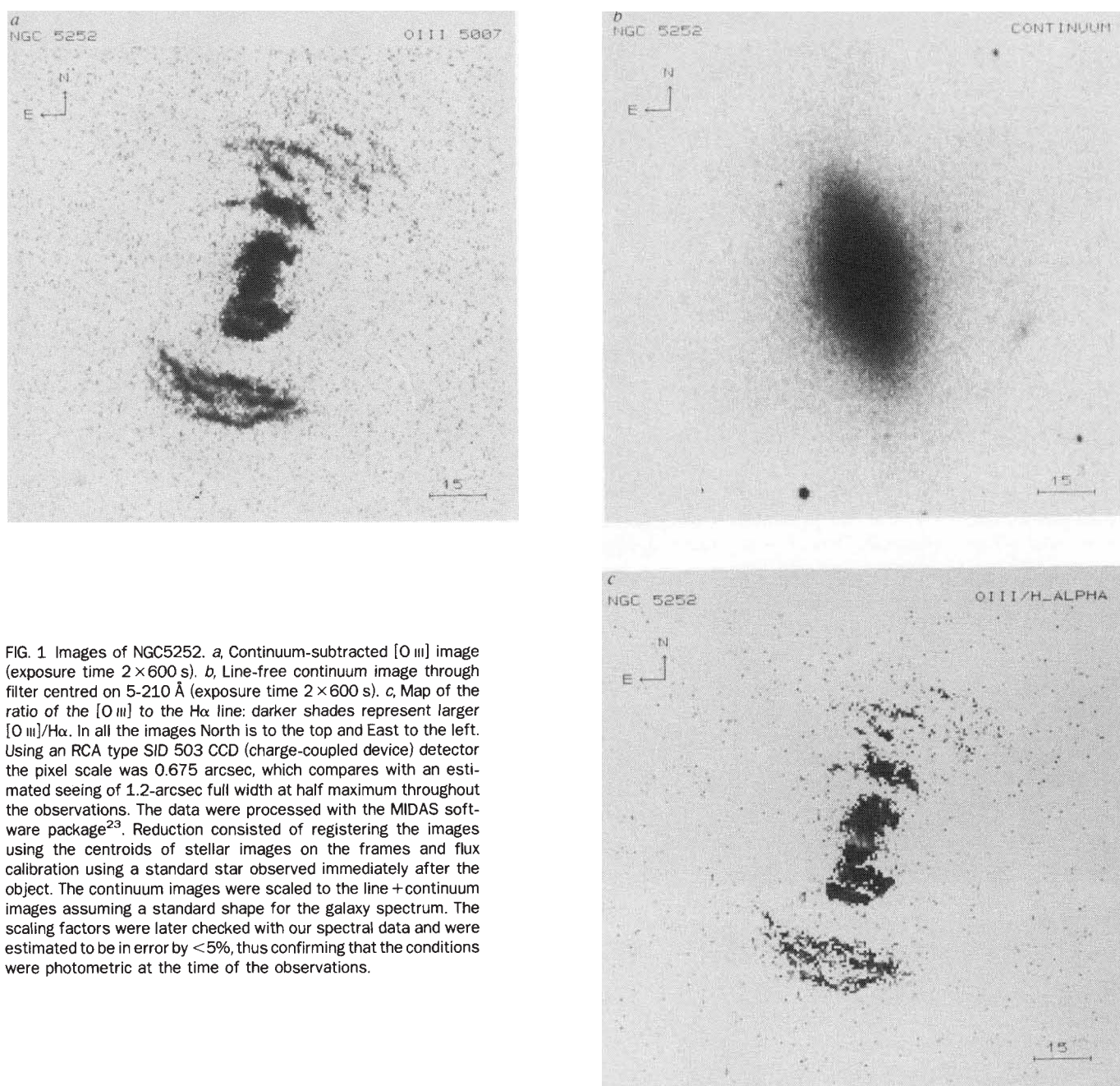


FIG. 1 Images of NGC5252. *a*, Continuum-subtracted [O III] image (exposure time $2 \times 600 \text{ s}$). *b*, Line-free continuum image through filter centred on $5-210 \text{ \AA}$ (exposure time $2 \times 600 \text{ s}$). *c*, Map of the ratio of the [O III] to the $\text{H}\alpha$ line: darker shades represent larger [O III]/ $\text{H}\alpha$. In all the images North is to the top and East to the left. Using an RCA type SID 503 CCD (charge-coupled device) detector the pixel scale was 0.675 arcsec , which compares with an estimated seeing of 1.2-arcsec full width at half maximum throughout the observations. The data were processed with the MIDAS software package²³. Reduction consisted of registering the images using the centroids of stellar images on the frames and flux calibration using a standard star observed immediately after the object. The continuum images were scaled to the line + continuum images assuming a standard shape for the galaxy spectrum. The scaling factors were later checked with our spectral data and were estimated to be in error by $< 5\%$, thus confirming that the conditions were photometric at the time of the observations.

that attempt to unify radio galaxies with quasars²⁰. Therefore, disk obscuration should be considered as a viable alternative to relativistic beaming as the radiation collimation mechanism in the powerful radio sources.

In general, our narrow-band images of NGC5252 support the idea that anisotropy/orientation effects play an important role in determining the appearances of active galaxies. □

Received 23 June; accepted 25 August 1989.

1. Lawrence, A. & Elvis, M. *Astrophys. J.* **256**, 410–426 (1982).
2. Antonucci, R. R. J. & Miller, J. S. *Astrophys. J.* **297**, 621–623 (1985).
3. Unger, S. W. et al. *Mon. Not. R. astr. Soc.* **228**, 671–679 (1987).
4. Wilson, A. S., Ward, M. J. & Haniff, C. A. *Astrophys. J.* **334**, 121–129 (1988).
5. van Breugel, W., Miley, G., Heckman, T., Butcher, H. & Bridle, A. *Astrophys. J.* **290**, 496–516 (1985).
6. Tadhunter, C. N. et al. *Mon. Not. R. astr. Soc.* **235**, 403–423 (1988).
7. Netzer, H. *Mon. Not. R. astr. Soc.* **216**, 63–78 (1985).
8. Orr, M. J. L. & Browne, I. W. A. *Mon. Not. R. astr. Soc.* **200**, 1067–1080 (1982).
9. Jackson, N., Browne, I. W. A., Murphy, D. W. & Saikia, D. J. *Nature* **338**, 485–487 (1989).
10. Robinson, A., Binette, L., Fosbury, R. A. E. & Tadhunter, C. N. *Mon. Not. R. astr. Soc.* **227**, 97–114 (1987).
11. Gottesman, S. T. & Harwarden, T. G. *Mon. Not. R. astr. Soc.* **219**, 759–775 (1986).
12. Pérez, E., González-Delgado, R. M., Tadhunter, C. N. & Tsvetanov, Z. *Mon. Not. R. astr. Soc.* (in the press).
13. Meurs, E. J. A. *Astrophys. & Space Sci.* **140**, 353–358 (1988).
14. Cheng, F. Z., Danese, L., de Zotti, G. & Fanceschini, A. *Mon. Not. R. astr. Soc.* **212**, 857–871 (1985).
15. Véron-Cetty, M.-P. *Astr. Astrophys. Suppl.* **58**, 665.
16. Stockton, A. & McKenty, J. W. *Astrophys. J.* **316**, 584–596 (1987).
17. Tadhunter, C. N. thesis, Univ. Sussex (1986).
18. Baum, S. A., Heckman, T. M., Bridle, A., van Breugel, W. J. M. & Miley, G. K. *Astrophys. J. Suppl.* **68**, 643–714 (1989).
19. Baum, S. A. & Heckman, T. M. *Astrophys. J.* (in the press).
20. Barthel, P. D. *Astrophys. J.* **336**, 606–611 (1989).
21. Nilson, P. *Uppsala General Catalogue of Galaxies Series V: A*, vol. 1 (Roy. Soc. Sci., Uppsala, 1973).
22. Edelson, R. A. *Astrophys. J.* **313**, 651–661 (1987).
23. *The MIDAS Users' Guide* ESO Operating Manual No. 1 (European Southern Observatory, 1988).

ACKNOWLEDGEMENTS. We thank B. Fosbury and E. Meurs for useful discussions. This work is based on observations collected at the European Southern Observatory, La Silla. C.T. was supported by European Space Agency External Fellowship. Z.T. is on leave from the Bulgarian Academy of Sciences.

Strike-slip faulting of ridged plains near Valles Marineris, Mars

Richard A. Schultz

Geodynamics Branch, Code 621, NASA Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

THE surface of Mars shows abundant evidence of extensional and contractional deformation in the form of normal faults, grabens and wrinkle ridges^{1–3}. In contrast, strike-slip faults have been considered to be extremely rare or absent on Mars^{4,5}. This view is based on the lack of significant lateral offset of pre-existing markers such as craters across suspected faults. Thus, Mars is believed to have evolved without plate tectonics or other large-scale lateral motions of the lithosphere⁶. However, careful study of Viking Orbiter images is revealing that strike-slip faults may be present on Mars⁷. Here I identify and document several well preserved examples of martian strike-slip faults and examine their relationships to wrinkle-ridges. The strike-slip faulting predates or overlaps periods of wrinkle-ridge growth south-east of Valles Marineris, and some wrinkle ridges may have nucleated and grown as a result of strike-slip displacements along the echelon fault arrays. Lateral displacements of several kilometres inferred along these arrays may be related to tectonism in Tharsis.

The inferred strike-slip faults deform moderately cratered plains of presumed volcanic origin south-east of Valles Marineris in Coprates⁸ (Fig. 1). Wrinkle ridges of north-south strike also deform plains materials throughout the area (Fig. 2). The strike-slip faults are defined by a series of echelon linear structures that bound polygonal or rhombohedral plateaus located within their stepovers (Fig. 3). These arrays of echelon structures and plateaus are oriented obliquely (~60°) to the overall trend of the wrinkle ridges. The arrays of echelon structures and plateaus are morphologically distinct from typical wrinkle ridges on the

Moon and Mars^{9,10}. The plateaus do not seem to result from disruption and lateral offset of pre-existing wrinkle-ridge topography because the ridges do not match up across the echelon structures. Further, the location of plateaus within stepovers argues against simple offset because not all of the echelon arrays that have plateaus are associated with wrinkle ridges or other high topography outside the stepover (Fig. 3a). Thus, the presence of polygonal plateaus within the stepovers of echelon structures suggests that the plateaus formed as a result of mechanical interaction between the faults. Local high topography and basins are typically produced within stepovers along echelon fault arrays. For example, extensional basins can occur within stepovers along overlapping spreading centres^{11,12} and extensional stepovers along strike-slip faults¹³; however, uplifts or basins are not generally associated with stepovers between dip-slip faults¹⁴. High topography can occur within contractional stepovers along strike-slip faults. Well-known examples include the East Bay Hills and Ocotillo Badlands uplifts along the San Andreas fault in California and horsts along the Hope strike-slip fault in New Zealand^{15–17}. The rhombohedral shape of certain plateaus, such as the large composite one in Fig. 3a, is consistent with the geometry of rock fracture in a strike-slip stepover (see ref. 14). Hence, the linear echelon structures in Coprates are probably strike-slip faults and the plateaus correspond to uplifts within contractional stepovers. Previous workers^{18–20,7} have interpreted echelon morphology within wrinkle ridges or other linear zones as evidence for strike-slip faulting. Here I document echelon geometry between wrinkle edges and note the scale independence of deformation associated with echelon geometries.

The echelon linear structures and plateaus on Mars are similar geometrically, and probably mechanically, to strike-slip faults and push-up ranges on the Earth (Fig. 4). The correlation between overlap and separation suggests that push-up length (overlap in Fig. 4) is related to propagation of faults in echelon arrays²¹ rather than being directly proportional to fault offset (see ref. 13). The minimum horizontal shortening associated with martian push-up ranges can be estimated by assuming that the plateaus are fault-bounded rigid blocks. Horizontal shortening is given by $S = 2(h/\tan d)$ for a push-up range bounded by two reverse dip-slip faults, where h is the push-up relief and d is the fault dip (30° for thrust faults). Plateau relief obtained from shadow-length measurements ranges from 200 m (composite push-up in Fig. 3a) to 780 m (superimposed ridges in Fig. 3b), with uncertainties of ~10–50%. Cumulative horizontal shortening along the north-west- and north-east-trending fault arrays (Figs 3a and b) is ~1.4 and 2.0 km, respectively. This kinematic estimate probably underestimates shortening because it assumes rigid materials with no internal deformation in the stepover region and neglects deformation within wrinkle ridges located near fault terminations. This relationship suggests that

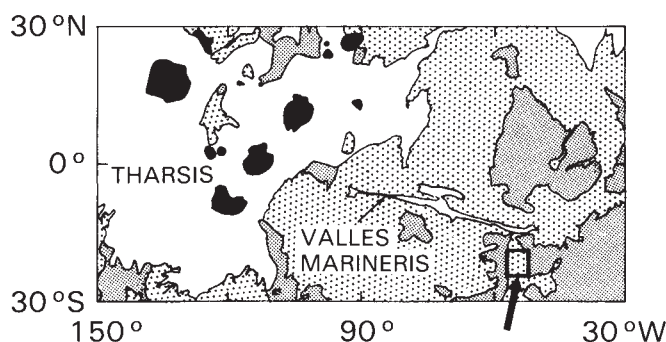


FIG. 1 Location map of western Mars indicating region of strike-slip faulting (inset marked by arrow). Dark shading, old (Noachian age) basement materials; light shading, intermediate age (Hesperian) plains materials; black, central volcanos; unpatterned, younger (Amazonian age) materials.

FIG. 2 On the left, a mosaic of medium resolution (220 m pixel^{-1}) Viking Orbiter images (610A27, 610A44, 610A46) showing north-trending wrinkle ridges (R) and oblique-trending strike-slip faults (arrows). 3a and 3b refer to enlargements in Fig. 3. Interpretation of relationships between echelon strike-slip faults and wrinkle ridges shown on the right. Orientations of maximum compressive principal stress inferred from left-lateral faults (σ_{3a}), N90W; from right-lateral faults (σ_{3b}), N80–85W.

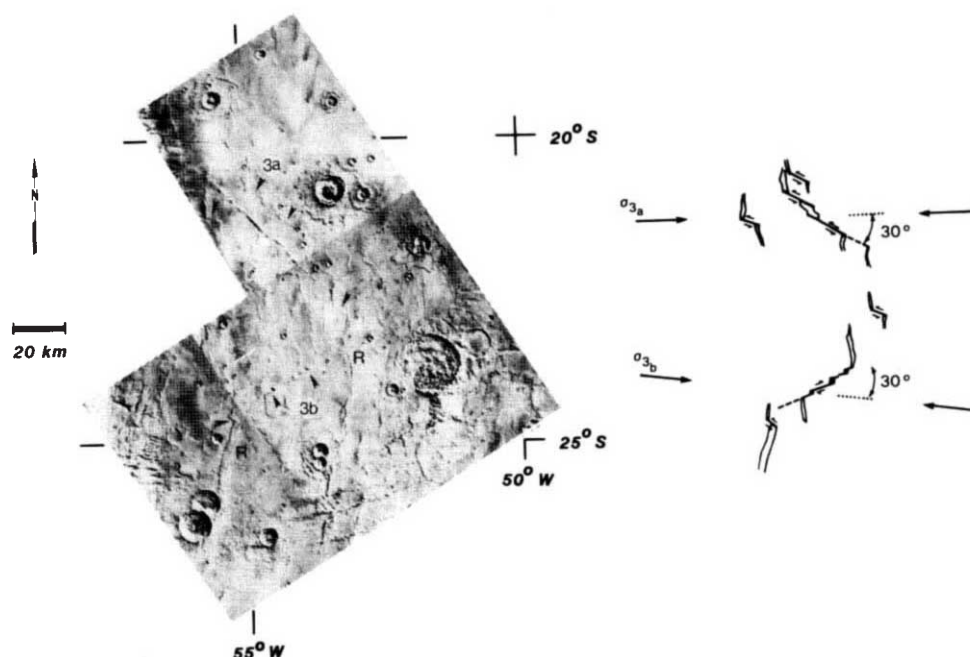
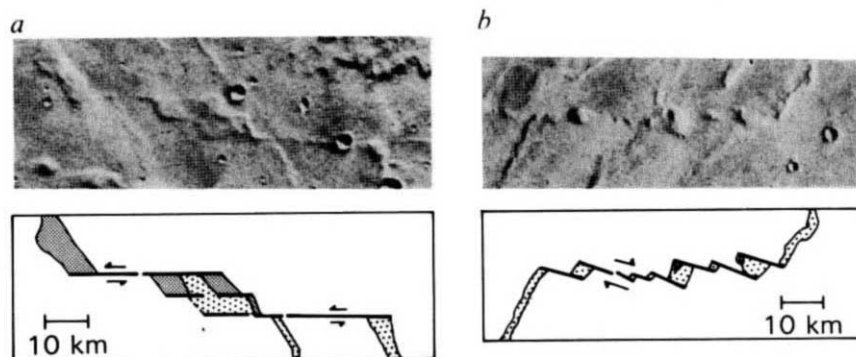


FIG. 3 Examples of martian echelon strike-slip faults. Plateaus and wrinkle ridges indicated by light shading; areas of high topography superimposed on other plateaus or ridges, dark shading. *a*, Right-stepping, left-lateral echelon faults. *b*, Left-stepping, right-lateral echelon faults.



the horizontal offset (shortening) along strike-slip faults in Coprates exceeds the vertical offset (push-up height) by at least a factor of two to three.

Determination of the sense of slip along echelon strike-slip faults is straightforward given the type of stepover and sense of step. For example, contractional stepovers occur for right steps along a left-lateral fault or left steps along a right-lateral fault. The north-west-trending strike-slip faults (Fig. 3a) step consistently to the right, implying left-lateral slip along the echelon array. The north-east-trending faults to the south (Fig. 3b) step consistently to the left, implying right-lateral slip along that array. Secondary extensional structures such as end-cracks, joints and normal faults or contractional structures such as stylolites, folds and reverse dip-slip faults can grow near strike-slip fault terminations as a result of fault slip^{22,23}, and the locations of these structures depend on the sense of slip. Consider a strike-slip fault in the map view located at the origin of a local xy coordinate system along the x -axis. For a right-lateral fault, large horizontal compressive stresses build up in the first and third quadrants, promoting nucleation and growth of folds and reverse dip-slip faults. Similarly, compressive stresses are increased in the second and fourth quadrants of left-lateral faults. The right-lateral fault arrays in Fig. 3b are associated with wrinkle ridges in the first and third quadrants, and left-lateral faults in Fig. 3a join wrinkle ridges in second and fourth quadrants. Wrinkle ridges are notably absent in the opposing, extensional quadrants of these faults. These relationships suggest that these wrinkle ridges may have nucleated in response to slip along the strike-slip faults and propagated away from

the fault-termination regions. High-angle (but not orthogonal) intersections between wrinkle ridges and the strike-slip faults are consistent with this hypothesis. The lack of observed orthogonal fault-ridge intersections implies that the faults did not serve as passive transform faults. The location and orientation of wrinkle ridges near terminations of strike-slip faults provide independent evidence that wrinkle ridges represent contractional deformation such as folding or reverse dip-slip faulting. In some cases, small crenulations that extend away from some fault terminations are superimposed on larger ridges (Fig. 3b), perhaps implying multiple stages of contractional deformation and local uplift. Angular relations between the strike-slip faults and wrinkle ridges and the localized, discrete deformation represented by both sets of structures suggest that inhomogeneous strain gave rise to ridged-plains deformation, not translation and rotation of rigid fault-bounded blocks.

Geometric relationships (see, for example, ref. 22) between several strike-slip faults and wrinkle ridges suggest that strike-slip faulting predated growth of at least some wrinkle ridges. In a few cases, however, apparent superposition of fault-termination wrinkle ridges on much larger ridges implies the reverse sequence. No examples were found that would suggest strike-slip faulting that clearly postdated wrinkle-ridge formation. Thus, it seems that strike-slip faulting predated and overlapped the episodes of wrinkle-ridge growth. Strike-slip and wrinkle-ridge deformation in Coprates are early-to-middle Hesperian in age ($\sim 3\text{--}3.5 \times 10^9$ yr BP, ref. 24).

The remote stress state associated with the martian strike-slip faults is inferred from their orientations and sense of slip. Using

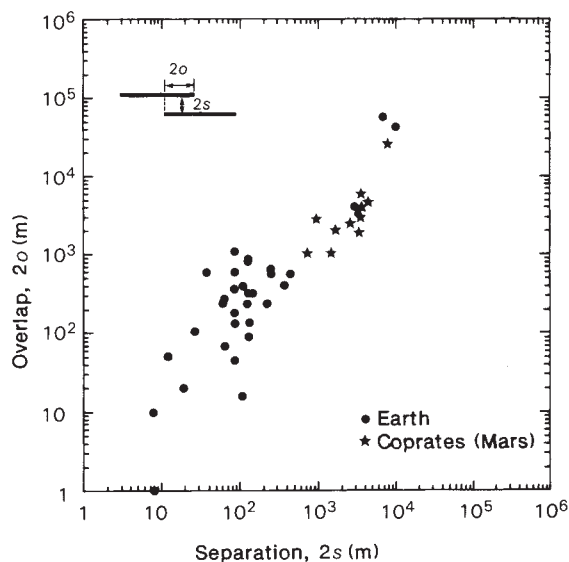


FIG. 4 Plot of overlap against separation of echelon strike-slip faults from Earth and Mars. Inset defines geometric parameters. Measurements of martian fault geometry from Fig. 2; terrestrial faults from ref. 21. Filled symbols, contractional stepovers. The geometry of echelon strike-slip faults on Mars is comparable to that of the terrestrial examples.

typical values of fault friction²⁵⁻²⁷ and a Coulomb slip condition, the direction of maximum horizontal compressive principal stress σ_3 (tension positive) would be oriented 30° to the faults, or approximately east-west (Fig. 2). Orientations of maximum compressive stress associated with wrinkle-ridge formation obtained by assuming that σ_3 is orientated normal to the overall trend of the ridges²⁸⁻³⁰ are very similar to those inferred from the strike-slip faults, and within $\sim 30^\circ$ of calculated Tharsis principal-stress trajectories²⁸⁻²⁹. It is likely that isostatic adjustments in the Tharsis region to the west²⁸ contributed to strike-slip faulting of Coprates ridged plains. The orientation of minimum compressive stress would, however, be vertical for wrinkle ridges but horizontal for strike-slip faults. Thus, strike-slip faults and wrinkle ridges can grow under the same remote stress state only if one set of structures changes the local stress state in such a way that the other structures can form. The geometry and relative timing of structures in Coprates indicate that some, but not all, wrinkle ridges nucleated and grew as a result of strike-slip faulting. Identification of the mechanisms for localizing strike-slip faulting in ridged plains and for relating along-strike variations of wrinkle-ridge morphology to ridge growth directions will help clarify the complex deformation history of martian ridged plains. □

20. Aubele, J. C. in *19th Conf. Lunar Planet. Sci.* XIX (abstract), 19-20 (1988).
21. Aydin, A. & Schultz, R. A. in *U.S. Geological Survey Open-File Rep.* 89-315, 47-66 (1989); *J. struct. Geol.* **315**, 47-66 (1989).
22. Rispoli, R. *Tectonophysics* **75**, 29-36 (1981).
23. Segall, P. & Pollard, D. D. *J. geophys. Res.* **85**, 4337-4350 (1980).
24. Tanaka, K. L. *J. geophys. Res.* **91**, E139-E158 (1986).
25. Byerlee, J. *Pure appl. Geophys.* **116**, 615-626 (1978).
26. Rice, J. R. *Pure appl. Geophys.* **121**, 443-475 (1983).
27. Li, V. C. in *Fracture Mechanics of Rock* (ed. Atkinson, B. K.) 351-428 (Academic, London, 1987).
28. Banerdt, W. B., Phillips, R. J., Sleep, N. H. & Saunders, R. S. *J. geophys. Res.* **87**, 9723-9733 (1982).
29. Sleep, N. H. & Phillips, R. J. *J. geophys. Res.* **90**, 4469-4489 (1985).
30. Watters, T. R. & Maxwell, T. A. *J. geophys. Res.* **91**, 8113-8125 (1986).

ACKNOWLEDGEMENTS. I thank R. Forsythe, D. Harding and T. Watters for comments on the manuscript and A. Aydin for insightful discussions. R.A.S. was supported by a Resident Research Associateship through the U.S. National Academy of Sciences/National Research Council.

Is ozone destroyed during the Antarctic winter in the presence of polar stratospheric clouds?

Giorgio Fiocco*, Walter D. Komhyr† & Daniele Fuà*‡

* University La Sapienza, Physics Department, 00185 Rome, Italy

† NOAA, Air Resources Laboratory, Boulder, Colorado 80303, USA

‡ Consiglio Nazionale Ricerche, 00044 Frascati, Italy

SINCE the discovery of the springtime Antarctic ozone depletion¹⁻² a great deal of attention has been given to processes related to aerosol formation and heterogeneous chemistry at low temperatures. Polar stratospheric cloud (PSCs),³ are frequent features of the southern winter atmosphere, and may provide a site for heterogeneous reactions that lead to ozone destruction^{4,5}. Models have been proposed⁶⁻⁸ for the formation of PSCs based on the condensation of HNO_3 (see ref. 9 for a review). An optical radar (lidar) has been operated at Amundsen Scott (South Pole), Antarctica, since the austral summer 1987-88. Observations made during the 1988 polar night show the presence of PSCs. Here we report ozonesonde measurements made quasi-simultaneously at the South Pole which indicate sharp minima of the ozone concentration in the vicinity of the PSCs. Although definitive information is unavailable for unambiguous interpretation of the results, the data may be viewed as evidence either for the role of dynamics in transporting air of different composition in conditions of substantial stability, or for processes leading to ozone destruction during the polar winter; the latter may include heterogeneous chemical reactions taking place in the absence of photolysis.

The lidar consists of a Nd YAG laser (532 nm) with a pulse repetition rate of 4 Hz, a Cassegrain receiving telescope of 0.5-m diameter and a two-channel photomultiplier detection system. A 1-MHz/12-bit digitizer inputs the data into a microcomputer that performs preliminary integration, storage and transmission of the data to the South Pole communications centre. The complete data set is stored on tape and retrieved when the station is no longer weather-bound. Meanwhile, samples of the data are transmitted to Rome by satellite.

As operated during the winter 1988, the instrument was capable of observing clouds and aerosols through the lower stratosphere. A more detailed description of the system and of the results will be given elsewhere. Lidar operation was not continuous during the first-year trial period; data were obtained, generally twice daily, during the periods 23-25 May, 21 June-12 July and 20 August-15 October. Drifts in the threshold of the sampling system and in the power emitted by the laser prevented data acquisition at altitudes ≥ 18 km.

Ozone and air-temperature profiles were obtained¹⁰ using electrochemical-concentration-cell (ECC) sondes (telemetry systems on balloons) launched at the South Pole on a weekly basis during January-August, but more frequently in September. In winter, standard meteorological radiosondes are launched daily around 22:30 GMT by weather station personnel.

Received 2 June; accepted 21 August 1989.

1. Plescia, J. B. & Saunders, R. S. *J. geophys. Res.* **87**, 9775-9791 (1982).
2. Solomon, S. C. & Head, J. W. *J. geophys. Res.* **87**, 9755-9774 (1982).
3. Chicarro, A. F., Schultz, P. H. & Masson, P. *Icarus* **63**, 153-174 (1985).
4. Wise, D. U., Golombek, M. P. & McGill, G. E. *Icarus* **38**, 456-472 (1979).
5. Golombek, M. P. *J. geophys. Res.* **90**, 3065-3074 (1985).
6. Solomon, S. C. *Geophys. Res. Lett.* **5**, 461-464 (1978).
7. Forsythe, R. D. & Zimbelman, J. R. *Nature* **336**, 143-146 (1988).
8. Blasius, K. R., Cutts, J. A., Guest, J. E. & Masursky, H. *J. geophys. Res.* **82**, 4067-4091 (1977).
9. Sharpton, V. L. & Head, J. W. *Proc. 18th Lunar Planet. Sci. Conf.* 307-317 (Cambridge University Press, 1988).
10. Watters, T. R. *J. geophys. Res.* **93**, 10,236-10,254 (1988).
11. Pollard, D. D. & Aydin, A. *J. geophys. Res.* **89**, 10,017-10,028 (1984).
12. Semper, J.-C. & Macdonald, K. C. *Tectonics* **5**, 151-163 (1986).
13. Aydin, A. & Nur, A. *Tectonics* **1**, 91-105 (1982).
14. Aydin, A. & Nur, A. in *Strike-Slip Deformation, Basin Formation, and Sedimentation, Soc. Econ. Paleont. Mineral. Spec. publ.* 37 35-44 (1985).
15. Aydin, A. & Page, B. M. *Geol. Soc. Am. Bull.* **95**, 1303-1317 (1984).
16. Sharp, R. V. & Clark, M. M. *U.S. geol. Surv., Prof. pap.* **787**, 131-140 (1972).
17. Freund, R. *N.Z. geol. Surv. Bull.* **86**, 1-48 (1971).
18. Tija, H. D. *Geol. Soc. Am. Bull.* **81**, 3095-3100 (1970).
19. Raitala, J. *Moon Planet.* **25**, 105-112 (1980).

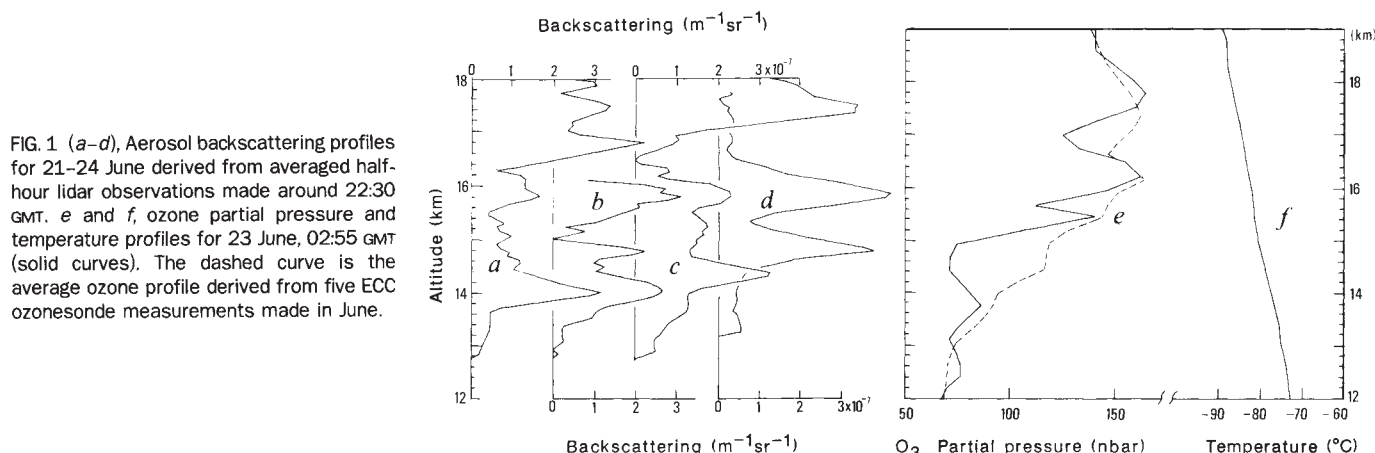


FIG. 1 (a-d), Aerosol backscattering profiles for 21-24 June derived from averaged half-hour lidar observations made around 22:30 GMT. e and f, ozone partial pressure and temperature profiles for 23 June, 02:55 GMT (solid curves). The dashed curve is the average ozone profile derived from five ECC ozonesonde measurements made in June.

An important feature of the lidar data is the frequent appearance of PSCs when temperatures are ≤ 200 K. (PSCs were observed at all times during late May-early August.) From the lidar echoes it is possible to infer the volume backscattering cross-section of the aerosols by assuming that in certain height regions no aerosols are present. Backscattering values are of the order of $10^{-7} \text{ m}^{-1} \text{ sr}^{-1}$, reaching a maximum of $6 \times 10^{-7} \text{ m}^{-1} \text{ sr}^{-1}$ at 14 km on 25 June. There is a considerable day-to-day variability in the location and intensity of the stratifications.

A perusal of the 1988 ozone-partial-pressure profiles¹¹ shows the general yearly trend, with the ozone hole phenomenon becoming apparent in early September. We focus our attention on one of several minima which appear in mid winter in otherwise smooth profiles. The 'ozone hole' is consistently evident: There are sharp features on all ozone profiles taken between 11 June and 22 July, at altitudes between 14 and 18 km, where temperatures are < 190 K. After 22 July the features disappear, to appear again around 1 September in smoother forms that can be seen as late as 19 October, which is the last day when temperatures are below the threshold for PSC formation.

A comparison of the aerosol and ozone profiles indicates that the ozone minima approximately coincide with the location of the PSC. The correspondence between aerosol layers and ozone minima is shown in Fig. 1 which depicts four half-hour-average aerosol profiles for 21-24 June (all taken around 22:30 GMT) and ozone and temperature profiles for 23 June, 02:55 GMT. (For reference, Fig. 1 also shows an average ozone profile derived from measurements made in June.) The ozone minima occur around 14.5, 15.6 and 16.9 km. Less than a day later (curve c), the intense stratifications around 14.4, 15.8 and 17.4 km were seen by lidar.

The columnar ozone depletion integrated through the height interval of the minima in the ozone profiles shown in Fig. 1 is of the order of $10^{21} \text{ molecule m}^{-2}$. Ozone minima and PSCs are present in all data sets obtained during June-July. Their spatial overlap seems to improve for lidar and ozone observations made closer in time. On 29 June, 22:43 GMT, aerosol stratifications occur at 15.8 km and 17.0 km and on the same day at 02:29 GMT a definite ozone minimum appears at 15.6 km. On 12 July at 22:55 GMT aerosol layers appear around 14.5 km and at 15-16 km, and on 13 July at 08:17 GMT ozone minima occur at 15.5 km and at 16.0 km.

The records indicate the disappearance of the ozone-reduction features during August; the lidar, however, was inoperative during July 13-August 19. When lidar operation resumed, PSCs could still be observed. In late August they maintained some of the sharp features evident in the first part of winter but during September such features tended to smooth out. This may be due to the onset of turbulent mixing as well as to the progressive hydration of the aerosol particles.

It is now generally accepted that the presence of PSCs in Antarctica induces a change in the equilibrium of atmospheric

trace gas species that, because of photolysis during daylight, induces the catalytic cycle responsible for the ozone hole. During the polar night, however, no evidence has been found of possible chemical destruction of ozone.

The simultaneous presence of the ozone and aerosol features may be ascribed to different mechanisms, as follows. (1) Air masses with low ozone contents and large amounts of condensable matter, possibly water, reach the South Pole stratosphere, structured in thin layers (1-km thick). No chemical interaction between constituents is implied: the ozone depletion is related to the origin of the air masses. (2) Air masses carrying PSCs pass in and out of darkness through sunlit regions where photochemical ozone destruction may occur. Thus, ozone may also be destroyed in Antarctica during winter and not only during springtime. (3) If the PSCs and HNO_3 condensates then ozone loss can occur in the formation of these clouds; NO_2 reacts with O_3 to yield N_2O_5 (this may occur in the absence of sunlight) and N_2O_5 is then converted to give HNO_3 . Although this is a possible process on a small scale, it is unlikely to occur at the magnitude implied by the measurements. (4) Ozone is destroyed by reaction on PSCs.

Regarding (1) and (2) above, the geopotential surfaces indicate that a small wavenumber-1 prevailed in Antarctica during the month of June and that the centre of the polar vortex was displaced from the South Pole. There is no obvious evidence, however, that trajectories passing over the South Pole traversed regions outside the polar night. The maintenance of the sharp features observed in the ozone and PSC profiles for about one month requires either a continuous supply of thinly stratified air or, if it is the same air that goes around unmixed, a very slow vertical diffusion process (that is, modest amounts of turbulence, probably in the presence of substantial shear).

Although weak ozone-aerosol correlations have been noticed in the past at mid-latitudes in the aftermath of volcanic eruptions, sharply correlated features such as we report here have not, to our knowledge, been observed before. $\square$

Received 21 February; accepted 25 August 1989.

1. Chubachi, S. *Mem. Natl. Inst. Polar Res. Jap., Spec. Iss.* **34**, 13-19 (1984).
2. Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207-210 (1985).
3. McCormick, M. P., Steele, H. M., Hamill, P., Chu, P. W. & Swisler, T. J. *J. Atmos. Sci.* **39**, 1387-1397 (1982).
4. Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. J. *Nature* **321**, 755-758 (1986).
5. McElroy, M. B., Salawitch, R. J., Wofsy, S. C. & Logan, J. A. *Nature* **321**, 759-762 (1986).
6. Cruzen, P. & Arnold, F. *Nature* **324**, 651-655 (1987).
7. Toon, O. B., Hamill, P., Turco, R. P. & Pinto, J. *Geophys. Res. Lett.* **13**, 1284-1288 (1986).
8. Hamill, P., Toon, O. B. & Turco, R. P. *Geophys. Res. Lett.* **13**, 1288-1291 (1986).
9. Solomon, S. *Rev. Geophys.* **26**, 131-148 (1988).
10. Komhyr, W. D. *Ann. Geophys.* **25**, 203-210 (1969).
11. Komhyr, W. D., Lathrop, J. A., Poston, R. W. & Mullen, T. D. *National Oceanic and Atmospheric Administration (Air Resources Laboratory), Data Report ERL ARL-18* (U.S. Department of Commerce, Silver Spring, 1989).

ACKNOWLEDGEMENTS. We thank R. Poston and T. Mullen for the lidar and ozonesonde observations and A. Adriani, M. Cacciani, G. Di Sarra, J. Breckenridge, C. Storey, and P. Di Girolamo for their help. We also thank J. DeLuise, P. Crutzen, P. Hamill and S. Solomon for discussions. We acknowledge the support of the Antarctic programmes of Italy and the United States and the NSF.

The role of cloud microphysics in high-cloud feedback effects on climate change

C. M. R. Platt

CSIRO, Division of Atmospheric Research, Station Street, Aspendale, Victoria 3195, Australia

THE effect of clouds on the radiation balance of the Earth (cloud forcing) is such that, on average, they cause the mean surface temperature to decrease. When the climate system is perturbed, the cloud forcing may change and this change will exert a feedback on the system. Climate models predict that clouds will have an overall positive feedback effect, that is, their presence will tend to enhance any surface-temperature changes, as in the greenhouse effect for example. The models preferentially increase the amounts of high clouds and their optical depths and decrease the amount of low clouds. Here I look at the temperature dependence of the infrared emittance of high clouds using data from a ground-based lidar/radiometer and *in situ* aircraft measurements of cloud properties. I conclude that the infrared emittances of high clouds are less sensitive to changes in temperature than predicted by climate models, chiefly because the microphysics of cirrus clouds modifies the relationship between cloud optical depth and cloud ice/liquid water path, in ways not accounted for in the models.

The annual-global-mean radiative temperature $\bar{T}_s$ is controlled by the balance of the incoming solar radiation that is absorbed by the planet and the terrestrial infrared radiation, the intensity of which is proportional to $\bar{T}_s^4$. The value of $\bar{T}_s$ is ~ 255 K or -18°C , which would be the Earth's mean surface temperature ($\bar{T}_g$) without any atmosphere. $\bar{T}_g$ is actually much greater ($\sim 12^\circ\text{C}$) because atmospheric trace gases, such as CO_2 and H_2O , are strong absorbers of infrared radiation over the whole terrestrial infrared emission spectrum, except for an 8–13 μm 'window' which allows surface radiation to escape to space. The absorbing gases re-emit to space high in the atmosphere to give an effective radiating temperature of $\bar{T}_s$. At the same time downward emission to the Earth increases the surface value well above $\bar{T}_s$ —the so-called greenhouse effect.

The value of $\bar{T}_g$ is further modified by clouds, which have a twofold effect. First, as they have a higher solar reflectance than most terrestrial surfaces, they reduce the radiation absorbed at the surface, thus reducing $\bar{T}_g$. Second, clouds absorb rather than scatter radiation at all infrared wavelengths, thus blocking emitted radiation from the surface in the atmospheric window, and re-emitting the radiation at cloud temperature, causing a reduction in planetary radiation to space and an increase in downwards radiation, thus increasing $\bar{T}_g$. The net magnitudes of the solar and infrared effects are such that low clouds cause a surface cooling, but high clouds result in surface warming^{1,2}. Averaged over all cloud types and amounts, there is a surface cooling and $\bar{T}_g$ is reduced by $\sim 18^\circ\text{C}$ from the cloudless value of $\sim 30^\circ\text{C}$ ^{3–5}. Thus, on average, clouds seem to reduce the equilibrium warming that is due to trace gases by a significant amount. From

these conclusions, however, we cannot predict what will happen when the climate is perturbed by an increase in CO_2 , for example, because different cloud types may not change their optical properties equally.

Incorporation of cloud properties into global climate models (GCMs) is still rather crude so that prediction of the effects of clouds on climate change is only approximate. Some GCMs do now carry interactive-cloud schemes so that clouds can respond to climatic perturbations^{6–9}, but only cloud amount can vary. Recently, however, variable cloud liquid-water content, linked to cloud solar reflectance and infrared emittance, has been incorporated^{10,11}.

The above GCMs have consistently shown that an effective doubling of CO_2 increases the high-cloud amount but decreases the lower-cloud amount. Both these changes will cause a further increase in $\bar{T}_g$, that is, a positive feedback. Calculations¹⁰ also show that the cloud liquid-water content increases everywhere. Although there is still a net planetary warming, $\bar{T}_g$ actually decreases¹². Roeckner^{10,11} performed a 20-year simulation using a GCM with coupled mixed-layer ocean and sea-ice models. The governing equations were solved at three vertical levels (850, 550 and 250 hPa) for a control model and for a model perturbed by 2% increase in the solar constant—roughly equivalent to a CO_2 doubling. They extracted changes in cloud quantities such as cloud liquid-water content (calculated from moisture continuity) and infrared emittances.

Here I compare the simulation^{10,11} results at 250 hPa with two sets of results based on observations of cirrus clouds that have recently become available. These data were taken at a wide range of cirrus temperatures, so that I could calculate the temperature derivatives. One set of data comprises particle size distributions of cirrus clouds measured *in situ* from an aircraft at different atmospheric levels and temperatures and in a number of cirrus clouds¹³. The data show that the cloud infrared absorption coefficient, the ice/water content and the cloud-particle mode radius all increase with temperature^{14,15}. The other set of data comprises remote observations on cirrus with a surface lidar and passive infrared radiometer¹⁶. These data include cloud infrared vertical-beam emittance, cloud depth and infrared absorption coefficient. The temperature variation of the absorption coefficient is similar to that found in the aircraft measurements. Some preliminary examinations¹⁷ of feedback have been done using the combined data sets.

To compare model and observational results the above observed quantities must be related to the variables of the model. For example, the measured cloud vertical-beam emittance ϵ_a (ref. 16) is the radiance in a narrow cone of angles in a direction perpendicular to the cloud divided by the blackbody radiance at cloud temperature and in the same cone. The diffuse-flux emittance ϵ in ref. 10, however, refers to the total flux emitted at all angles from the cloud as measured by a horizontal detector, divided by the total cloud blackbody flux¹⁸. Now ϵ_a can be written $\epsilon_a = 1 - e^{-\delta_a}$ where δ_a is the cloud absorption optical depth. It can be written as $\delta_a = K_a L = \sigma_a h$ where K_a is a mass absorption coefficient, L is the integrated vertical ice (or water) path, σ_a is the volume absorption coefficient, h is the cloud depth and I assume vertical homogeneity in the cloud. L is related

TABLE 1 Cirrus mean cloud-water content and infrared absorption parameters deduced from lidar/radiometer and cloud microphysical measurements

T ($^\circ\text{C}$)	h (km)	ϵ_a	δ_a	σ_a (km^{-1})	K_a ($\text{g}^{-1} \text{m}^2$)	L (g m^{-2})	dL/dT ($\text{gm}^{-2} \text{ } ^\circ\text{C}^{-1}$)	W (g m^{-3})
-70	1.30	0.085	0.089	0.065	0.133	0.67	0.062	5.1×10^{-4}
-60	1.78	0.130	0.139	0.073	0.084	1.65	0.154	9.3×10^{-4}
-50	2.27	0.215	0.242	0.107	0.051	4.75	0.440	2.1×10^{-3}
-40	2.76	0.310	0.371	0.134	0.032	11.6	1.08	4.2×10^{-3}
-30	2.67	0.425	0.553	0.207	0.019	28.4	2.660	1.1×10^{-2}

Data are deduced from measurements reported in refs 13–16.

TABLE 2 Comparison of emittance and water-path changes derived in this study with those predicted in ref. 10

	T (°C)	ΔT (°C)	ε	$\Delta \varepsilon$	L (g m ⁻²)	ΔL (g m ⁻²)
11° S*	-43	+7	0.42	0.18	—	5
11° S†	-43	+7	0.482	0.093	8.4	6.8
45° S*	-48	+5	0.42	0.24	—	4
45° S†	-48	+5	0.395	0.087	4.7	3.7

* Ref. 4.

† Data from lidar/radiometer and cloud microphysical measurements¹³⁻¹⁶.to ice/water content W by $L = Wh$ so that

$$K_a = \sigma_a / W \quad (1)$$

Now σ_a and W are proportional to the square and cube of the particle radius, respectively, and so¹⁷ $\sigma_a = 2W/\pi r_e$ or $K_a = 2/\pi r_e$ where r_e is the cloud-particle mode radius. Thus, σ_a depends on both W and r_e .

As the simulation results give the changes of emittance it is instructive to see how changes in ε with T relate to the other quantities. As ε is a function of ε_a

$$\frac{d\varepsilon}{dT} = \frac{d\varepsilon}{d\varepsilon_a} \frac{d\varepsilon_a}{dT} \quad (2)$$

or

$$\frac{d\varepsilon}{dT} = \frac{d\varepsilon}{d\varepsilon_a} (1 - \varepsilon_a) \left(K_a \frac{dL}{dT} + L \frac{dK_a}{dT} \right) \quad (3)$$

Various quantities derived from the observational data are shown in Table 1. Values of ε_a , h and σ_a were obtained from the lidar and radiometer studies¹⁶ and values of K_a were obtained from the cloud-particle spectra¹⁵. Values of σ_a previously calculated¹⁵ from cloud-particle spectra agreed well with the lidar/radiometer data¹⁵. Values of K_a from the cloud-particle spectra were extrapolated for temperatures $< -55^\circ\text{C}$. For consistency, values of W were calculated from equation (1), using the values of σ_a and K_a in Table 1. L was then calculated using the observed values of h .

Table 1 illustrates the large increases in W and L with temperature observed in high clouds. The mass absorption coefficient K_a decreases significantly, however, because the cloud-particle mode radius also increases, with the result that increases in σ_a (or δ_a) with temperature are much less than the increases in W (or L). (That is, the last term on the right-hand side of equation (3) is negative). These results imply that absorption (or extinction) must be parameterized in terms of r_e as well as W .

Changes in ε and L with temperature are compared with the model results¹⁰ at 250 hPa in Table 2. The temperatures shown in column 2 were taken from ref. 19. As the temperature changes ΔT quoted in ref. 10, were appreciable, comparison values of $\Delta \varepsilon$ and ΔL were obtained from plots of ε and L as a function of T (using the data in Table 1) rather than from equation (3). The initial values of L are not quoted in ref. 10.

There is reasonable agreement between the values of ΔL , but values of $\Delta \varepsilon$ obtained from the simulation seem to be too high by factors of two or three. Stephens's parameterization²⁰, used in ref. 10, is considered quite adequate for lower-boundary-layer clouds and some deeper water clouds, but does not take account of the large changes in r_e , and therefore in K_a , in cirrus clouds. The values of K_a used in ref. 10 were apparently too high, implicitly neglecting the observed decrease in K_a with temperature. Somerville and Remer²¹ introduced the inclusion of liquid water into GCMs but they also made δ proportional to L only. My results show the necessity of allowing for r_e .

Apparently Roeckner's calculated values of ΔT that are due to cirrus cloud are too large, because the increase in cirrus

emittance is smaller than he predicted. As pointed out^{10,12}, the warming at cirrus altitudes will be counteracted by a net cooling at the surface because of a loss in solar radiation. According to Schlesinger¹², the overall radiative change is still positive, but my results indicate that this may not always be the case. $\square$

Received 9 February; accepted 15 August 1989.

- Stephens, G. L. & Webster, P. J. *J. Atmos. Sci.* **38**, 235-247 (1981).
- Charlock, T. P. *Tellus* **34**, 245-254 (1982).
- Cess, R. D. & Potter, G. R. *Tellus* **39**, 460-473 (1987).
- Ramanathan, V. et al. *Science* **243**, 57-63 (1989).
- Manabe, S. & Wetherald, R. T. *J. Atmos. Sci.* **24**, 241-259 (1967).
- Hansen, J. A. et al. in *Climate Processes and Climate Sensitivity*, *Geophys. Monogr. Ser.* **29**, 130-163 (1984).
- Wilson, C. A. & Mitchell, J. F. B. *J. geophys. Res.* **92**, 13,315-13,343 (1987).
- Wetherald, R. T. & Manabe, S. *J. Atmos. Sci.* **45**, 1397-1415 (1988).
- Schlesinger, M. E. & Mitchell, J. F. B. *Rev. Geophys.* **25**, 760-798 (1987).
- Roeckner, E. *Atmos. Res.* **21**, 293-303 (1988).
- Roeckner, E. et al. *Nature* **329**, 138-140 (1987).
- Schlesinger, M. E. *Nature* **335**, 303-304 (1988).
- Heymsfield, A. J. *J. Atmos. Sci.* **34**, 367-391 (1977).
- Heymsfield, A. J. & Platt, C. M. R. *J. Atmos. Sci.* **41**, 846-855 (1984).
- Platt, C. M. R. in *Proc. Int. Radiation Symp., Perugia, 1984*, 163-166 (Deepak, Hampton, Virginia, 1984).
- Platt, C. M. R., Dilley, A. C. & Scott, J. C. *J. Atmos. Sci.* **44**, 729-747 (1987).
- Platt, C. M. R. & Harshvardhan, J. *geophys. Res.* **93**, 11,051-11,058 (1988).
- Platt, C. M. R. & Stephens, G. L. *J. Atmos. Sci.* **37**, 2314-2322 (1980).
- McClatchey, R. A., Fenn, R. W., Selby, J. E. A., Volz, F. E. & Garing, J. S. *Optical Properties of the Atmosphere* (3rd edn, Rep. No. AFCRL-72-8497 (Air Force Cambridge Research Laboratories, Massachusetts, 1972).
- Stephens, G. L. *J. Atmos. Sci.* **35**, 2123-2132 (1978).
- Somerville, R. C. J. & Remer, L. A. *J. geophys. Res.* **89**, 9668-9674 (1984).

Reconstruction of tree-line vegetation response to long-term climate change

Serge Payette, Louise Filion, Ann Delwaide & Christian Bégin

Centre d'Études Nordiques, Université Laval, Sainte-Foy, Québec, Canada, G1K 7P4

KNOWLEDGE of the vegetation response to climate change is necessary to assess and predict realistic ecosystem development in the anticipated, CO₂-induced warmer world, particularly at high latitudes where greater warming is expected¹⁻³. Reconstruction of vegetation development over the past 1,000 years may be helpful in this respect, because this period was characterized by contrasting climatic conditions⁴⁻⁹. Here we report the reconstruction of wind-exposed, tree-line vegetation associated with long-term climate change in northern Canada, using tree-ring and growth-form analyses of spruce subfossils. Three major types of growth form within the exposed, but stable, lichen-spruce community successively predominated in response to climate forcing: high krummholz (dwarf spruce, <2-m high) with scarce small (>2-m high) trees (AD 1305-1435, cool period), trees (>2-3 m high) and high krummholz (AD 1435-1570, warm period) and low krummholz (≤ 50 cm) (little ice age to present: AD 1570 onwards, cold period and present climate, respectively). Whereas the expansion of a marginal lichen-spruce woodland climaxed during the late Middle Ages (AD 1435-1570), present development of a low-krummholz vegetation at these sites seems to be out of phase with the twentieth century warming. This suggests that ecosystem recovery to global warming is not straightforward, depending on the nature of vegetation structure present at the time climate change occurred. The implications of such ecosystem resilience for the detection and monitoring of the expected CO₂-induced warming is discussed, particularly for the climate-sensitive arctic and subarctic regions.

High-latitude tree-line sites are susceptible to changes in climate, where summer temperature¹⁰ and winter conditions¹¹⁻¹³ represent the main restrictive factors for tree growth. Vegetation responses to climate change at such marginal sites involve com-

plex interactions between the two closely related processes of sexual reproduction and phenotypic plasticity in tree species. In black spruce (*Picea mariana* (Mill.) BSP), the dominant tree species in subarctic Canada¹⁴, sexual reproduction occurs mostly in monopodial trees with normal stems, whereas krummholz are generally unable to produce a large number of cones and viable seeds¹⁵. The reaction of spruce to climate change may thus be different according to the type of dominant growth forms prevailing at the time the climate change occurred.

Old-growth tree-line vegetation that escaped fires during the past centuries in northern Québec is made of living and dead black-spruce stems of all ages^{16,17}. Depending on the microsite conditions, dead stems remain undecomposed for several hundred years, 400–600 yr being the average apparent residence time for logs found at the soil surface. This provides the opportunity to reconstruct the recent stand history from living and dead material of wind-swept, old-growth coniferous sites in the Boniface River/Busk Lake area of northern Québec (57° 50' N, 76° W). The area has a rolling, low-plateau topography made of small hills (the so-called upland sites: 150–200 m) protruding the lowlands (115–125 m). Present vegetation at all of the study sites is typically lowland and upland tundra—scattered or dense low krummholz, with dwarf birch (*Betula glandulosa* Michx.) and fruticose lichens. Small forest stands are located in snowy (80–150 cm of snow), protected sites. At upland sites, the stunted spruce (<30–50 cm high) have very narrow tree rings which are difficult to measure. The average snow depth rarely exceeds 30 cm, as suggested by mean height of woody vegetation. In contrast to the living vegetation in upland sites, leaning dead spruces are abundant at the soil surface and include several types of shrubby and arboreal growth forms. Tree subfossils in wind-swept lowland sites, however, are very infrequent. Spruce stumps were recovered from mires and, during a period of exceptionally low water level, from lakes.

Four sets of data were sampled. The first two sets included subfossil trees and the third set comprised subfossil high krummholz. Together these formed a mixed subfossil population that lay on the ground of exposed, low-krummholz communities situated on five small hilltops (Fig. 1). The fourth set, composed of large subfossil tree stumps recovered from lowland mires and lakes within a 50-km² area, was sampled for a comparative tree-ring and growth-form analysis with upland subfossils. Growth form, total height and winter damage (indicated by multiple-stem and branch leaders) of dead spruces were evaluated for all specimens found on the small hills presently under severely exposed snowless conditions (Fig. 1). A Henson micrometer (0.01-mm precision) was used for tree-ring measurements of finely sanded stem sections. Standard tree-ring techniques¹⁸ were applied only to stem disks without reaction wood¹⁹ (eccentric tree ring development in the direction of stems tilted towards the ground). Although missing or incomplete tree rings were common, cross-dating of spruce wood was facilitated by the available light-ring (tree rings without latewood cells or showing reduced latewood-cell development) chronology²⁰ supplemented by new records of light-ring years, especially between AD 1305–1560 (A.D., L.F. and S.P., unpublished data). Because an unknown proportion of outer rings was removed from dead stems by weathering and decomposition, reported dates of stem mortality may be in error by 10–20 yr. Four tree-ring chronologies based on spruce growth forms were deduced: upland-tree chronology from disks above the inferred snow-air interface (AST (above-snow tree), AD 1372–1799, including 34 radii from 21 samples), upland-tree chronology from disks below the inferred snow-air interface (BST (below-snow tree), AD 1350–1803, including 28 radii from 16 samples), upland-krummholz chronology from disks below the inferred snow-air interface (BSK (below-snow krummholz), AD 1305–1811, including 23 radii from 15 samples) and lowland-tree chronology (LT, AD 1425–1789, including 13 radii from 7 samples). In Fig. 2 we present curves of the mean width of tree

rings of these spruce growth forms, which give direct evidence of past climate change.

Cross-dating of subfossil stem parts used in the tree chronologies indicate that normal trees developed at both upland and lowland sites between AD 1435 and 1570 (Fig. 2, AST and LT time series). Periods of faster growth recorded in the AST chronology around AD 1435–1450, 1495–1525 and 1550–1570 coincided with active upright-stem development above the critical snow-air interface, whereas only small stems developed around AD 1375, 1585 and 1740. Based on the frequency of stem development above the inferred snow-air interface, upland sites were occupied by lichen-spruce communities made up mostly of high krummholz with scarce small trees between AD 1305 and 1435. Marginal woodlands dominated by high krummholz and larger trees expanded between AD 1435 and 1570. According to the morphological features of subfossil spruces, the average snow-air interface was in the range 50–80 cm above the ground surface before AD 1570, after which it decreased steadily to its modern level of 30 cm. As stem development above the inferred snow-air interface also declined after AD 1570, clearly shown by the tree-ring patterns in all tree chronologies, stem dieback intensified particularly during the

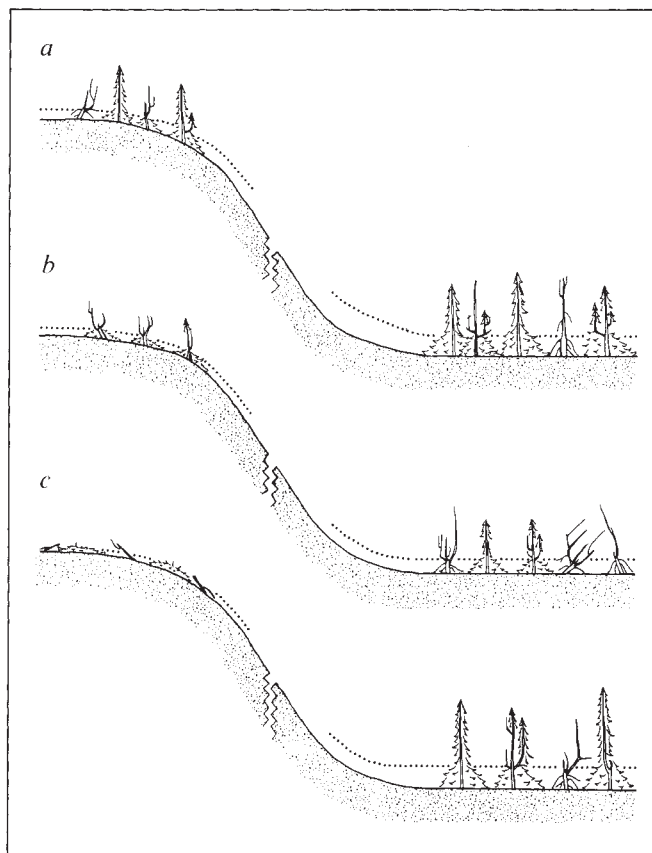


FIG. 1 Vegetation reconstruction in exposed tree-line sites from a, late Middle Ages to b, little ice age and c, the present, in comparison with spruce development in snow-protected forest sites during the same period (ref. 16 and C.B., unpublished data). Presence or absence of vertical stem development above the snow cover (dotted line) constitute the best indication of present and past climatic conditions in cold subarctic environments. Sampling was stratified according to position above and below the inferred snow-air interface. Three data sets were used in the reconstruction. The first two sets associated with subfossil trees (in black) were composed of tree-stem sections or disks sampled respectively below and above the inferred snow-air interface shown in c, as determined by morphological traits such as abrupt changes in stem diameter and multiple stem leaders. The third set was made of subfossil krummholz (in black) sections all located below the inferred snow-air interface. Vegetation reconstruction was inferred from cross-dated subfossil spruces, which show distinct growth-form patterns (krummholz and tree) through time.

17th century (Fig. 2, AST and LT time series). From AD 1570–1580 onwards, a period corresponding to the onset and full development of the little ice age, the rate of stem dieback became considerably higher than production of new above-snow stems. Only snow-protected parts of tree and high krummholz survived at these sites until the 18th and 19th centuries, when new growth was in the form of small, horizontal spruce layers (Fig. 2, BST and BSK time series). Because sustained dieback proceeded continuously during this period, marginal lichen-spruce woodlands that developed during the AD 1435–1570 period progressively changed towards exposed, snowless, low-krummholz communities. Most of the creeping spruces which dominate present upland vegetation are survivors, through the process of layering, of 16th century spruces.

Tree-ring patterns of the four chronologies were closely associated with the temporal development of the spruce growth forms (Fig. 2). Highly significant correlations (0.01 level) were found between all time series (r ranging from 0.61 to 0.74, except between the BSK and LT chronologies where $r = 0.46$). The strongest associations are between the tree time series, which indicate quite similar growth conditions associated with the climate and growth habit. The lowest mean ring width and standard deviation of ring widths were from the BSK time series (0.18 ± 0.06 mm), which expressed the slow development of high krummholz with their foliage and growth meristems mostly concentrated less than 1 m above the ground. Faster growth and higher variability of ring width in the AST (0.21 ± 0.10 mm).

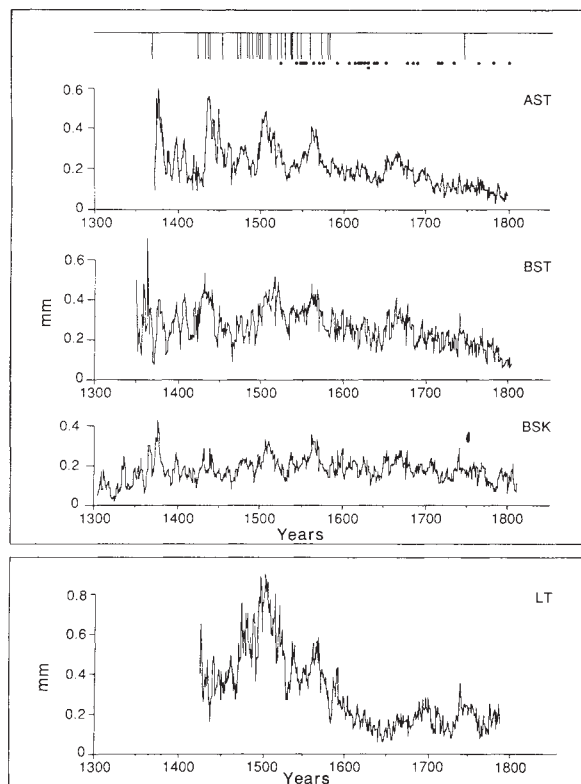


FIG. 2 Time series of ring widths from upland and lowland sites at the tree-line. Upland-tree chronologies: AST (disks sampled above the snow-air interface) and BST (disks sampled below the snow-air interface). Upland-krummholz chronology: BSK (disks sampled below the snow-air interface). Lowland-tree chronology: LT, shown separately for comparison. Vertical bar: tree-ring date of onset of individual stem development above the snow-air interface. Black dot: maximum tree-ring date (apparent age +20 yr to account for the removal of outer rings by weathering and decomposition) of dieback of individual stems above the snow-air interface.

BST (0.26 ± 0.09 mm) and LT (0.31 ± 0.19 mm) chronologies reflect the irregular periods of stem growth and dieback. Time-series mean ring widths calculated for the two major growth periods, as defined by tree development (before AD 1570) and tree dieback (after AD 1570), indicated relatively rapid growth before AD 1570 among tree chronologies (AST: 0.28 ± 0.10 mm before AD 1570, 0.16 ± 0.06 mm after AD 1570; BST: 0.31 ± 0.09 mm before AD 1570, 0.22 ± 0.07 mm after AD 1570; LT: 0.48 ± 0.16 mm before AD 1570, 0.18 ± 0.07 mm after AD 1570), whereas krummholz growth (BSK) remained fairly constant throughout (0.18 ± 0.07 mm and 0.18 ± 0.05 mm, respectively).

Such direct responses of upland lichen-spruce communities to climate change may help to predict future directions of vegetational change in the climate-sensitive Arctic and subarctic areas. A major feature to consider is the site sensitivity and the associated vegetation resilience to climate disturbance. Comparison of developmental sequences among strategic sites positioned along a wind-exposed toposequence indicate that, before the onset of the little ice age, both exposed (this study) and protected sites¹⁶ (C.B., unpublished data, as shown in Fig. 1) responded positively to warming by favouring tree expansion and development of open woodlands, whereas severely exposed sites (composed of living and dead spruces showing the same prostrate, low-krummholz growth form) remained fairly stable from the late Middle Ages to the present (Fig. 1). Thus, differential vegetation responses to climate change only occurred during this century at the two former sites. Although protected lichen-spruce sites showed a shift from high krummholz to forest from the little ice age to 20th century, most of the exhausted creeping krummholz in exposed lichen-spruce sites have been unable to react positively to present warming. This situation may be explained by a marked difference in vegetation resilience between the lichen-spruce ecosystems. High krummholz located in protected sites were able to catch drifting snow and maintain the snow cover above the critical level for spruce survival during the little ice age. At exposed sites, the inability of creeping spruces to control the snow-drift environment was a key factor enhancing snowless conditions. Such conditions prevented seed production and seedling establishment in upland sites, whereas in protected sites large seed bearers were already present when recent warming started¹⁶. Whether the shift from snow-rich to snowless conditions in upland sites was associated with an exacerbated wind-disturbance regime caused by death of large pre-little-ice-age spruces unable to trap blowing snow, or to a major change in winter climate, or to both influences, is a question that deserves consideration.

Our data suggest that vegetation responses to global warming are not as straightforward as one may expect, as it is necessary to sort out the magnitude of climate change and the importance of autogenic control involved in ecosystem resilience. Vegetation resilience to environmental change, as reported here along a wind exposure gradient, constitutes an important ecosystem attribute. The amplitude and duration of recent warming were probably not large enough to compensate for the detrimental effects of the long-lasting cooling of the little ice age on exposed sites. Accordingly, recovery of the stand to its state before the little ice age would require a much milder winter climate in order to significantly increase winter survival of buds and needles, a necessary step to achieve stem development, cone production and spruce regeneration. Although recent climatic data indicate sustained global warming²¹ during this century, no conclusive evidence of a positive vegetation response to such warming has yet been identified at these exposed tree-line sites. A reliable test for assessment of substantial global warming associated with rising atmospheric carbon dioxide at the arctic-subarctic transition, that is, warming that could overcome current wind and snow effects, would thus be from direct field evidence of vertical growth-form development leading to a type of ecosystem at least similar to that of the AD 1435–1570 period. □

Received 2 May; accepted 23 August 1989.

1. Hansen, J. *et al. Science* **213**, 957–966 (1981).
2. Manabe, S. & Stouffer, R. J. *J. geophys. Res.* **85**, 5529–5550 (1980).
3. Manabe, S. & Wetherald, R. T. *J. Atmos. Sci.* **37**, 99–118 (1980).
4. LaMarche, V. C. Jr *J. Science* **183**, 1043–1048 (1974).
5. Kullman, L. *New Phytol.* **108**, 101–110 (1988).
6. Kullman, L. *J. Biogeogr.* **14**, 1–8 (1987).
7. Lamb, H. H. *Climate, Present, Past and Future* Vol. 2 (Methuen, London, 1977).
8. LeRoy Ladurie, E. *Histoire du Climat Depuis l'An Mil* (Flammarion, Paris, 1967).
9. Grove, J. M. *The Little Ice Age* (Methuen, London, 1988).
10. Black, R. A. & Bliss, L. C. *Ecol. Monogr.* **50**, 331–354 (1980).
11. Hadley, J. L. & Smith, W. K. *Ecology* **67**, 12–19 (1986).
12. Wardle, P. *Arct. alp. Res.* **13**, 419–423 (1981).
13. Tranquillini, W. *Physiological Ecology of the Alpine Timberline* (Springer, Berlin, 1979).
14. Hosie, P. C. *Native Trees of Canada* (Supplies and Services Canada, Ottawa, 1980).
15. Payette, S. & Gagnon, R. *Nature* **313**, 570–572 (1985).
16. Payette, S., Filion, L., Gauthier, L. & Boutin, Y. *Nature* **315**, 135–138 (1985).
17. Filion, L., Payette, S. & Gauthier, L. *Géogr. phys. Quat.* **34**, 221–226 (1985).
18. Fritts, H. C. *Tree Rings and Climate* (Academic, New York, 1976).
19. Scurfield, G. *Science* **179**, 647–655 (1973).
20. Filion, L., Payette, S., Gauthier, L. & Boutin, Y. *Quat. Res.* **26**, 272–279 (1986).
21. Jones, P. D. *et al. Nature* **332**, 790 (1988).

ACKNOWLEDGEMENTS. We thank K. Gajewski for commenting on the manuscript. We acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada, the Ministère de l'éducation du Québec (FCAR Program), and the Geological Survey of Canada.

Earthquake recurrence and risk assessment in circum-Pacific seismic gaps

Wayne Thatcher

US Geological Survey, Menlo Park, California 94025, USA

THE development of the concept of seismic gaps, regions of low earthquake activity where large events are expected, has been one of the notable achievements of seismology and plate tectonics. Its application to long-term earthquake hazard assessment continues to be an active field of seismological research. Here I have surveyed well documented case histories of repeated rupture of the same segment of circum-Pacific plate boundary and characterized their general features. I find that variability in fault slip and spatial extent of great earthquakes rupturing the same plate boundary segment is typical rather than exceptional but sequences of major events fill identified seismic gaps with remarkable order. Earthquakes are concentrated late in the seismic cycle and occur with increasing size and magnitude. Furthermore, earthquake rupture starts near zones of concentrated moment release, suggesting that high-slip regions control the timing of recurrent events. The absence of major earthquakes early in the seismic cycle indicates a more complex behaviour for lower-slip regions, which may explain the observed cycle-to-cycle diversity of gap-filling sequences.

The seismic gap concept originated from the work of Fedotov¹ and Mogi², who examined the relation between seismicity and the occurrence of great earthquakes on the margin of the north-west Pacific. They noted that over a time interval of about 30 years, the spatial distribution of background seismic activity defined a long linear belt interrupted by a number of conspicuous gaps. The seismicity belt was comprised of the aftershock zones of individual great earthquakes, which outlined adjoining regions but did not significantly overlap. In the areas of low activity, the seismic gaps, there had been no large or great earthquakes during the preceding decade or more. Historical records indicated that great events had occurred in the gaps during the late 19th or early 20th centuries, however, and several of the gaps did rupture in the years 1952–1963. Fedotov and Mogi recognized the forecasting implications of this observation and suggested that large events should be expected in the remaining gaps. Between 1968 and 1973 there were major earthquakes³ in four of these gaps.

The discoveries of plate tectonics immediately supplied a simple rationale for these empirical inferences⁴. The long linear seismic belt defined the boundary between the Pacific and Eurasian lithospheric plates. The gaps represented unruptured segments in locked frictional contact where relative plate motion was building up elastic stress that would ultimately be relieved in great earthquakes. These conclusions quickly led to a recognition that the seismic-gap concept could be applied to make qualitative assessments of long-term seismic hazard on major plate boundaries throughout the world^{3–6}. These applications have been remarkably successful: up to 1988, 13 identified circum-Pacific seismic gaps have subsequently ruptured in large or great earthquakes⁷.

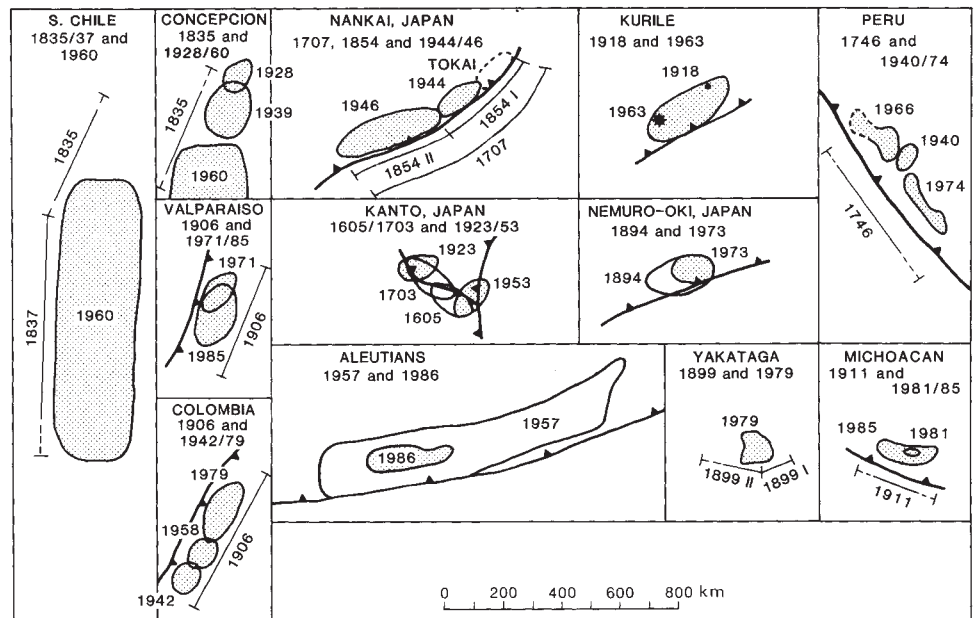
During the past ten years, research has increasingly been directed towards quantifying the timing of long-term earthquake predictions and integrating these estimates into a probabilistic framework^{8–13}. At the same time, new work has highlighted several unresolved issues which limit the confidence that can be placed in the newer quantitative assessments. Gap-filling events have differed in size from those expected^{14–17}, some observations suggest very irregular recurrence^{18,19} and many recurrence estimates critically depend on the assumed behaviour of plausible but idealized models^{8–13,20,21}. These difficulties raise questions concerning the degree to which long-term hazard can be quantified and the applicability of simple conceptual models to earthquake recurrence. To clarify these issues and the origins of observed recurrence behaviour, I have examined observations of repeated rupture of specific segments of subduction zones in the circum-Pacific region. These are summarized for the two most recent cycles of earthquake strain release in Fig. 1, and further details of these and earlier events are contained in refs 8, 14–17, 22–26.

An important feature of Fig. 1 is the complete absence of repeated events with similar source characteristics. Commonly, a single long rupture is succeeded by several gap-filling events in the next episode of earthquake strain release (Valparaíso, Colombia, Peru, Michoacán). Sometimes this involves encroachment of rupture into or from an adjacent segment (Southern Chile, Concepción). In several instances, not all of the plate boundary segment defined by preceding events has ruptured (Nankai and Kanto) and for several recent shocks (1979 Yakataga, 1986 Aleutians) only a small fraction of the seismic gap has yet been filled. Although the 1918 and 1963 Kurile earthquakes are generally similar, their mainshock epicentres are separated by 200 km.

A second notable feature of Fig. 1 is the timing, source size and magnitude pattern of gap-filling episodes that involve several events. These characteristics are summarized in Fig. 2, which shows that each event in the sequence invariably occurs in the last half of the seismic cycle, the majority being concentrated in the last 20% of the cycle. Furthermore, events show a strong tendency to increase in magnitude and rupture length toward the end of the cycle. The cycle duration in each region is defined by the largest events in the two most recent strain-release episodes. In four of the regions, times between the preceding great earthquakes are comparable to cycle durations defined in this way^{8,17,22}. For Peru the average interval is 194 yr for the past 2 cycles, for Concepción 98 yr (4 cycles), for Valparaíso 85 yr (4 cycles), and for Nankai 112 yr (4 cycles). In the three other regions (Colombia, Kanto, Michoacán), the seismic history is too poorly known to independently confirm my choice of cycle duration^{15,23,26}.

In many quantitative assessments of long-term risk the size of gap-filling events is not directly addressed and it is tacitly assumed that the expected earthquake will have the same features as the most recent preceding event. Previous studies have shown, however, that in several regions gaps have only been partially filled²⁷, events were smaller than expected¹⁴ or several shocks occurred when only a single event was anticipated^{15–17}. The larger dataset shown in Fig. 1 shows that it has

FIG. 1 Map views of extent of gap-filling earthquake ruptures on 12 segments of circum-Pacific plate boundary. North is towards the top of each frame. The trench axis, the surface trace of the convergent-plate boundary, is shown by a thick solid line with teeth on the over-riding lithospheric plate. Extent of rupture for each event is given by aftershock zone, tsunami source area or distribution of seismic intensities. In each region the extent of the most recent event or sequence of events is indicated by stippling. For some of the earlier events the area of rupture is not well determined and source extent is indicated by straight line segments, dashed where uncertain.



been difficult to forecast the character of gap-filling events because cycle-to-cycle diversity is seen to be typical rather than exceptional. At the same time, the concentration of sequences of events late in the seismic cycle can be of practical value in anticipating future activity, as the partial rupture of an identified gap may signal the imminence of more (and larger) events. Conversely, if major portions of identified seismic gaps have ruptured in one or more great events, the results summarized in Fig. 2 suggest that smaller shocks will not fill in the remainder of the gap until late in the subsequent cycle of seismic-strain release (for example, the Tokai segment in Nankai region, Fig. 1).

Historical documents or geological methods are increasingly being applied to determine the times between previous earthquakes^{28,29}, providing a firmer perspective on the significance and level of hazard implied by the present repose intervals in unruptured gaps. Whereas probabilistic calculations usually assume that inter-event times are tightly clustered about a mean recurrence interval and some observations support this assumption³⁰, geological investigations on the San Andreas fault in southern California have recently uncovered evidence of very irregular recurrence^{18,19}. The temporal pattern of gap-filling earthquake sequences shown in Fig. 2 has several similarities with the southern California observations. There, inter-event times range from as short as 45 yr to more than 300 yr, but the ten dated events occur in clusters of two or three shocks that span intervals of ≤ 100 yr and are separated from adjacent clusters by 300–400 yr. Except for the historical earthquake of 1857, the magnitude and source extent of the California events is unknown. Furthermore, whereas the circum-Pacific sequences usually occur in spatially distinct rupture zones that do not significantly overlap, the California chronology is determined at a single site. Nonetheless, the temporal concentration of events is strikingly similar, suggesting an underlying order in the recurrence of the San Andreas events that mimics the circum-Pacific sequences and implies a common mechanism.

When detailed earthquake chronologies are unavailable but the long-term fault slip or relative plate-motion rate is known, fault slippage in the most recent great earthquake has been used to estimate the time required to reaccumulate that slippage and generate the next great event (elastic-rebound model³¹ or time-predictable²⁰ model). Frequently the detailed slip distribution in individual events is unknown and average or local slip estimates are used in recurrence calculations^{8–13,20,21}. Although the general validity of the elastic-rebound model is widely acknowl-

Furthermore, whenever earthquake slip distributions have been well determined they invariably show considerable spatial heterogeneity, as Fig. 3 illustrates. Although it is difficult to edged and some data support the expectations of the time-predictable model^{20,21}, the detailed precision of such occurrence-time predictions is uncertain.

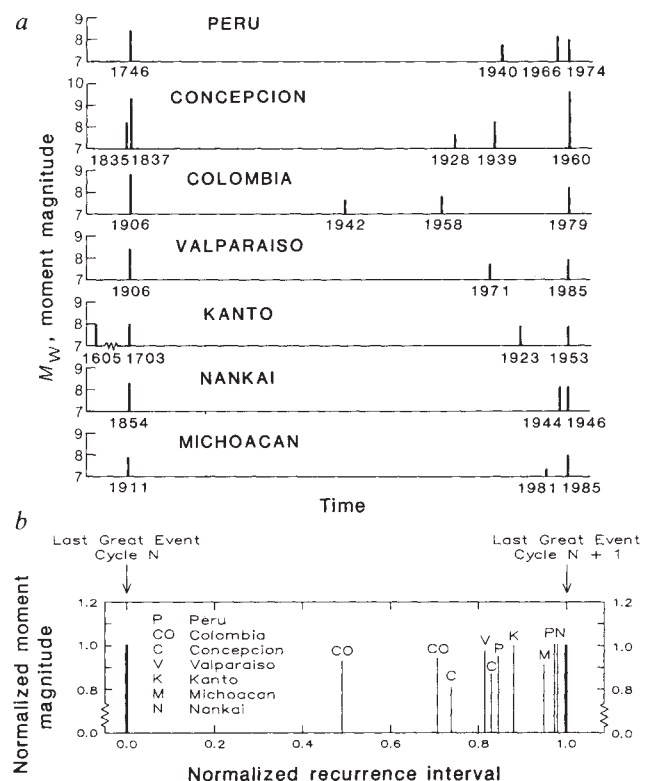


FIG. 2 Magnitude as a function of time for earthquakes from the seven regions in Fig. 1 where the most recent episode of seismic strain release has involved more than one great gap-filling event. The duration of the seismic cycle is defined by the time interval between the last event in each strain-release episode. *a*, Time-magnitude plots for each region, with duration of cycle normalized to the same measured length on the time axis of each plot. *b*, Data from *a* plotted on a single graph, with earthquake magnitude normalized by the last event of the most recent sequence in each region.

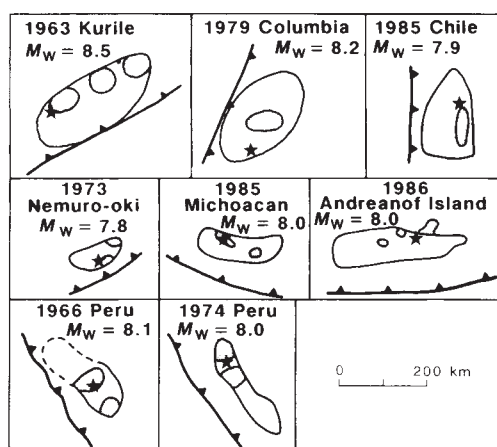


FIG. 3 Spatial distribution of seismic moment release for eight of the most recent great earthquakes shown in Fig. 1^{16,17,32-35}. Extent of rupture in each event is delineated by aftershock zone. Regions of concentrated moment release are indicated by stippling. Mainshock epicentre (the point of rupture initiation) is located by a star. North is towards the top of each figure and trench axis is given by heavy line with teeth on overthrust lithospheric plate.

estimate slip distributions directly by seismic methods, the spatial pattern of seismic moment release (the product of slip, fault area and elastic shear modulus) is an informative measure of slip that is being determined for an increasing number of large events. Moment-release distributions for eight of the recent great earthquakes considered earlier (Fig. 1) are plotted schematically in Fig. 3. These results demonstrate that seismic slip is typically nonuniform but also show an association between mainshock epicentre and regions of concentrated moment release. For all but one of the eight events (1979 Columbia) the point of rupture initiation lies either within or immediately adjacent to the high moment-release regions, which themselves comprise only a small fraction of the entire rupture plane. A larger data set of 23 events (W.T., manuscript in preparation) shows the same general features. This association suggests that earthquake recurrence is more likely to be controlled by the maximum rather than the average coseismic slip, leading to consistent underestimates of inter-event times if average values are used.

The observations reported here suggest a relation between earthquake slip distribution and recurrence behaviour but also demonstrate complexities that are not easily explained by simple models. The close proximity of the mainshock nucleation point to zones of concentrated moment release suggests that the high-slip regions determine the event times. The time and magnitude patterns of event sequences shown in Fig. 2 can be rationalized if the high-slip regions rupture according to the time-predictable model, with failure occurring at stress thresholds determined by slip maxima in the preceding cycle of strain release. To the degree that cycle duration is uniform these high-slip zones may persist from event to event and sustain the same amount of slippage in each cycle of seismic strain release.

The behaviour of regions with less slip is more problematic. Because slip rate is essentially constant along the strike of major plate-boundary rupture zones, the spatial heterogeneity of earthquake slip generates regions of major slip deficiency with the occurrence of every great event. If these slip-deficient regions behaved time predictably, recurrent events would be expected as soon as plate motion reaccumulated slip to the level of that released in the preceding great earthquake. The considerable nonuniformity of earthquake slip indicates that such events should occur relatively early in the seismic cycle, and the concentration of event sequences in the latter 20% of the seismic cycle (Fig. 2) is a strong argument against simple time-predictable behaviour in the slip-deficient regions. □

Received 28 June; accepted 21 August 1989.

1. Fedotov, S. A. *Trans. Inst. Fiz. Zemli Akad. Nauk SSSR* **36**, 66-93 (1965).
2. Mogi, K. *Bull. Earthquake Res. Inst. Tokyo Univ.* **46**, 175-203 (1968).
3. McCann, W. R., Nishenko, S. P., Sykes, L. R. & Krause, J. *Pure appl. Geophys.* **17**, 1082-1147 (1979).
4. Sykes, L. R. *J. geophys. Res.* **76**, 8021-8041 (1971).
5. Kelleher, J. *J. geophys. Res.* **77**, 2087-2103 (1972).
6. Kelleher, J., Sykes, L. R. & Oliver, J. *J. geophys. Res.* **78**, 2547-2585 (1973).
7. Nishenko, S. P. *Reinhold Encyclopedia of Geophysics* 260-268 (Van Nostrand, 1989).
8. Nishenko, S. P. *J. geophys. Res.* **90**, 3589-3615 (1985).
9. Lindh, A. G. *U.S. geol. Surv. Open File Rep.* **83-63**, 1-15 (1983).
10. Sykes, L. R. & Nishenko, S. P. *J. geophys. Res.* **89**, 5905-5927 (1984).
11. Nishenko, S. P. & Singh, S. K. *Bull. seism. Soc. Am.* **77**, 2095-2114 (1987).
12. Working Group on California Earthquakes Probabilities *U.S. Geol. Surv. Open File Rep.* **88-387** (1988).
13. Nishenko, S. P. & Jacob, K. *J. geophys. Res.* (in the press).
14. Abe, K. *J. Phys. Earth, Suppl.* **25**, 261-271 (1977).
15. Kanamori, H. & McNally, K. *Bull. seism. Soc. Am.* **72**, 1241-1254 (1982).
16. Beck, S. L. & Ruff, L. J. *J. geophys. Res.* **92**, 14,123-14,138 (1987).
17. Beck, S. L. & Ruff, L. J. *J. Phys. Earth planet. Inter.* (in the press).
18. Jacoby, G. Sheppard, P. & Sieh, K. *Science* **241**, 196-199 (1988).
19. Sieh, K., Stuiver, M. & Brillinger, D. *J. geophys. Res.* **94**, 603-623 (1989).
20. Shimazaki, K. & Nakata, T. *Geophys. Res. Lett.* **7**, 279-282 (1980).
21. Sykes, L. R. & Quittmeyer, R. C. in *Earthquake Prediction, An International Review* vol. 4 (eds Simpson, D. W. & Richards, P. G.) 217-247 (American Geophysical Union, Washington, DC, 1981).
22. Ando, M. *Tectonophysics* **27**, 119-140 (1975).
23. Matsuda, T., Ota, Y., Ando, M. & Yonekura, N. *Bull. geol. Soc. Am.* **89**, 1610-1618 (1978).
24. Engdahl, E. R., Kisslinger, C. & Billington, S. *J. geophys. Res.* (in the press).
25. McCann, W. R., Perez, O. J. & Sykes, L. R. *Science* **207**, 1309-1314 (1980).
26. Astiz, L., Kanamori, H. & Eissler, H. *Bull. seism. Soc. Am.* **77**, 1326-1346 (1987).
27. Ishibashi, K. in *Earthquake Prediction, An International Review* vol. 4 (eds Simpson, D. W. & Richards, P. G.) 297-332 (American Geophysical Union, Washington, DC, 1981).
28. Sieh, K. E. in *Earthquake Prediction, An International Review* vol. 4 (eds Simpson, D. W. & Richards, P. G.) 181-208 (American Geophysical Union, Washington, DC, 1981).
29. Sykes, L. R., Kisslinger, J. B., House, L., Davies, J. N. & Jacob, K. H. *Science* **210**, 1343-1345 (1980).
30. Nishenko, S. P. & Buland, R. *Bull. seism. Soc. Am.* **77**, 1382-1399 (1987).
31. Reid, H. F. *The California Earthquake of April 18, 1906, Report of the State Earthquake Investigation Commission* vol. 2, 16-28 (Carnegie Institution of Washington, Washington, DC, 1910).
32. Beck, S. L. & Ruff, L. J. *J. geophys. Res.* **89**, 9281-9291 (1984).
33. Choy, G. L. & Dewey, J. W. *J. geophys. Res.* **93**, 1103-1118 (1988).
34. Kikuchi, M. & Fukao, Y. *Tectonophysics* **144**, 231-247 (1987).
35. Houston, H. thesis, California Inst. Technol. (1986).

ACKNOWLEDGEMENTS. I thank H. Kanamori, A. G. Lindh, S. P. Nishenko and C. H. Scholz for discussions and J. C. Savage and R. S. Stein for their reviews.

Post-depositional stability of long-chain alkenones under contrasting redox conditions

F. G. Prahl*, G. J. de Lange†, M. Lyle‡
& M. A. Sparrow*

* College of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA

† Department of Geochemistry, Institute of Earth Sciences, University of Utrecht, Budapestlaan 4, PO Box 80.021, 3508TA Utrecht, The Netherlands

‡ Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York 10964, USA

PRYMNESIOPHYTE algae, which include the coccolithophorid species *Emiliania huxleyi*¹, are the recognized biological source of a series of long-chain (C_{37} , C_{38} , C_{39}), unsaturated methyl and ethyl ketones² widely observed in marine sediments³. Studies of *E. huxleyi* in culture have demonstrated that these biomarkers are attractive geochemical tools for palaeoceanographic study. Notably, unsaturation patterns within the alkenone series change regularly with growth temperature³⁻⁵ and the total alkenone abundance in the living plant cell is relatively constant, accounting for 5-10% of the total cellular organic carbon⁵. If these compounds are relatively well-preserved in sediments, profiles of alkenone unsaturation patterns and total alkenone concentration with depth in dated deep-sea cores⁶⁻⁸ provide a temporal record of sea surface temperatures and the productivity of an important group of marine phytoplankton³⁻⁵. Here we analyse the long-chain alkenone composition of sediment samples from above and below an oxidation front in an ungraded turbidite layer⁹ deposited 140 ± 12 kyr BP in the Madeira Abyssal Plain¹⁰, to evaluate the post-depositional stability of these biomarkers under contrasting redox conditions. The results demonstrate that >85% of the total amount

of these compounds is degraded over ~8 kyr as a consequence of diffusion-controlled oxidation¹⁰. Remarkably, such extensive degradative loss has little effect on the unsaturation pattern of the residual biomarker series. Thus, we find that long-chain alkenones provide reliable indicators of sea surface temperature in the ocean. The total abundance of these biomarkers in sediments, however, is controlled not only by prymnesiophyte productivity, but also by their degree of exposure to oxidative degradation in the sedimentary process.

Sediment cores from the Madeira Abyssal Plain (MAP), located in the north-east Atlantic Ocean, are characterized by a periodic sequence of thick (1–4 m) distal turbidites interbedded by thin (3–15 cm) pelagites^{10–12}. The main source of these organic-rich turbidite materials is the continental margin off north-west Africa located ~700 km to the south-east. Nineteen major turbidite layers of late Quaternary age have been identified in various cores from MAP and correlated stratigraphically over distances of 200 km in this deep-sea basin, demonstrating the large geographic scale of these sediment transport events^{11,13}. In general, the major turbidite layers observed in piston cores from MAP are ungraded silty clays¹⁴ consisting of 40–70% biogenic calcite (primarily coccoliths) with the remainder composed predominantly of quartz and clay minerals. Some of the turbidite layers contain a relatively high total-organic-carbon (TOC) content (~1%) compared with the interbedded pelagic intervals (~0.2%). The surface portion (15–45 cm) of these turbidites is often distinguished from the deeper portion by a colour transition and by a significant reduction in TOC (~80%) and calcium carbonate (~10%) content. These observations imply that the upper portion of these turbidites has undergone post-depositional oxidation resulting from the diffusion of oxygen into the surface of the deposit during intervals of slow pelagic sedimentation^{10,15}. The time required for the full evolution of these oxidation imprints on the geochemistry of these sediment records ranges from 2 to 15 kyr depending on the turbidite interval considered¹⁰.

The two sampling sites from MAP examined in this study (86P5: 32°2.50' N, 24°12.50' W; 86P25: 30°44.23' N, 25°22.45' W) are separated spatially by >100 km. One sample from the oxid-

ized layer and two samples from the unoxidized layer of the f-turbidite^{10,11} identified in each of these cores were taken for analysis of their bulk inorganic and organic carbon content and their long-chain alkenone composition and content. The total volume of sediment represented by the f-turbidite at MAP is estimated¹³ to be ~126 km³. The age of these sediments is estimated at 140 ± 12 kyr, based on biostratigraphic and oxygen isotope age determinations of the thin pelagic strata located above and below this turbidite interval¹⁰.

For the two sample sets taken from the f-turbidite^{10,11}, similar geochemical trends are noted at both bulk and molecular levels (Table 1). The concentrations (per g dry sediment) of TOC, calcium carbonate and total (C₃₇, C₃₈, C₃₉) alkenones are highest and nearly constant in the lower, unoxidized portion of the turbidite. Quite notably, the concentration for these three geochemical properties measured in the unoxidized intervals of the f-turbidite differs only slightly between these two widely separated sampling locations (86P5 and 86P25). These observations support previous evidence that MAP turbidites are distal and ungraded^{11,14}. By contrast, a pronounced difference is evident when the chemical composition of the oxidized and unoxidized intervals of this turbidite are compared. For both cores, the TOC, calcium carbonate and total alkenone content of the oxidized samples are lower than the corresponding unoxidized samples by consistent factors of 83%, 6% and ~88%, respectively. This observation argues that the diffusion-controlled oxidation process acted uniformly on the f-turbidite deposited throughout this basin.

A compositionally uniform series of long-chain alkenones is widely observed in marine sediments ranging in age from recent to Tertiary³. The predominance of these biomarkers of prymnesiophyte algae² in the total-solvent-extractable-lipid fraction of sediments^{3,16} argues that they are relatively more stable than most other lipids biosynthesized by marine plankton. Based on a close agreement between long-chain alkenone fluxes measured in sediment traps deployed at shallow depths for short periods of time and in underlying bottom sediments, Volkman *et al.*¹⁷ suggested that these biomarkers are transported conservatively from surface waters to the ocean bottom in the Peru Upwelling

TABLE 1 Geochemical data obtained from analyses of oxidized and unoxidized sediment intervals from the f-turbidite in two cores (86P5 and 86P25) from Madeira Abyssal Plain

Depth interval (cm)	Core 86P5			Core 86P25		
	oxidized 704–714	unoxidized 751–761	unoxidized 772–782	oxidized 780–789	unoxidized 855–868	unoxidized 905–916
Chemical parameter						
%Organic carbon*	0.14	0.89	0.91	0.18	0.97	0.99
%Calcium carbonate*	48.5	51.7	51.8	47.9	50.7	50.5
Total alkenones†						
ng per g dry	940 ± 62	10,300	9,000	1,240	8,620	8,800
µg per gC	685 ± 64	1,160	990	690	890	890
U ₃₇ ^{k'}	0.72 ± 0.01	0.71	0.68	0.72	0.71	0.70
SST (°C)	19.9 ± 0.2	19.6	19.0	19.9	19.8	19.4
K37/K38	1.3 ± 0.10	1.2	1.3	1.3	1.3	1.5
EE/K37	0.024	0.032	0.042	0.031	0.035	0.030

Piston cores (86P5: 32° 02.50' N, 24° 12.50' W; 86P25: 30° 44.23' N, 25° 22.45' W) were collected from 5,400-m water depth by the Dutch Geological Survey (RGD) in 1986 aboard RV *TYRO*.

* Total organic and inorganic carbon determinations were made by LECO carbon analyser using the wet oxidation technique described in ref. 23.

† Total long-chain (C₃₇, C₃₈, C₃₉) alkenones were extracted from wet sediment by soxhlet using an azeotropic mixture of toluene in methanol and isolated from the total extract by column chromatography on silica gel. The column chromatographic fraction containing these compounds was analysed by gas chromatography using a 30 m × 0.25 mm inner diameter DB-5 fused-silica column (J&W Scientific), splitless injection, temperature programming (75–130°C at 10°C min⁻¹, 130–300°C at 5°C min), hydrogen as a carrier gas (0.703 kg cm⁻² head pressure) and flame ionization detection. Individual compounds were quantified by an internal standard method using hexamethylbenzene as the internal standard. Further details of these methods are given in ref. 8. The unsaturation index U₃₇^{k'} {=(C_{37:2}/(C_{37:2} + C_{37:3}))} is defined by the concentration ratio of the di-unsaturated C₃₇ alkenone to the sum of the di- and tri-unsaturated C₃₇ alkenone. This index is related to the growth temperature T of the phytoplankton source by the equation: U₃₇^{k'} = 0.034 × T + 0.039. The parameters K37/K38 and EE/K37 are defined by the concentration ratio of total C₃₇:C₃₈ alkenones and of di-unsaturated C₃₆ ethyl alkenoate (EE) to total C₃₇ alkenones, respectively, measured in each mixture. Further description of these parameters is given in ref. 5.

Zone. In an open ocean environment, Prah1 and Muehlhausen¹⁸ compared the relative abundance of total alkenones to TOC measured in a sediment trap deployed at 3,900 m for >1 yr and in underlying surface sediments (4,400 m). Their results showed that these compounds were 2–3 times more reactive than TOC. Although the alkenones are selectively preserved relative to most other planktonic lipids, the latter observation implies that these biomarkers of prymnesiophyte algae are not a conserved property of sedimentary particles, at least in deep-water open-ocean environments.

Comparison of the total alkenone concentrations (per g C) measured in the oxidized and unoxidized portions of the f-turbidite (Table 1) yields an observation that agrees with the findings of the open-ocean sediment trap experiment. Total alkenones are not enriched relative to TOC in the oxidized portion of the turbidite as would be expected if these compounds were completely insensitive to biodegradation. Rather, these compounds are depleted to a somewhat greater extent than TOC (88% compared with 83%) by the diffusion-controlled oxidation process¹⁰ that acted on the surface layers of the f-turbidite.

Similar to TOC¹⁹, only a small fraction of the total quantity of long-chain alkenones biosynthesized in surface waters seems to survive transport through the water column and become preserved as a stable component of the sediment record. Emerson's model²⁰ for marine sediments suggests that the TOC burial flux in the deep sea depends on surface supply (productivity) as well as bottom-water oxygen content. Decreased availability of oxygen (and perhaps other oxidants such as nitrate) in pore waters should enhance the preservation efficiency of TOC in sediments. The magnitude of a change in TOC content noted down the core can, therefore, differ from that attributable to an absolute change in primary productivity if the bottom-water oxygen varies with time. The similar reactivity of the alkenones relative to TOC under contrasting redox conditions (Table 1) suggests that the stratigraphic record for the abundance of these biomarkers in sediments may also contain a significant, time-dependent preservational overprint.

Most palaeoceanographic studies are conducted with sediment cores from open-ocean environments. In these settings, the depth of zero oxygen in sediments is typically subsurface regardless of the magnitude of the surface-water productivity. Nonetheless, the depth of this boundary is not static but moves up or down in the underlying sediments in response to an increase or decrease in the supply of metabolizable organic carbon from above²¹. Hence, a change in productivity changes the thickness of the oxic layer in the sediment deposit and, therefore, the post-depositional exposure time of TOC and the alkenone biomarkers to oxygen. The results of this study do not address the sensitivity of these compounds to degradation under different oxygen exposures. Without a calibration of the behaviour of these biomarkers under different periods of oxygen exposure, stratigraphic records for alkenone concentration^{6–8} can provide at best only a semi-quantitative sense of how the prymnesiophyte productivity in overlying surface waters has varied at a given ocean location over time.

In culture, *E. huxleyi* exhibits unsaturation patterns within the alkenone series that change regularly with growth temperature^{3,5}. A plot of the index U_{37}^k (defined in Table 1) quantifies the degree of unsaturation in a given series as a function of growth temperature³. T follows a remarkably linear trend ($r^2 = 0.994$) defined by the equation $U_{37}^k = 0.034 \times T + 0.039$ (ref. 5). A test of this equation in the field has shown that it reliably predicts the unsaturation patterns measured in phytoplankton collected from ocean waters of known temperature⁴. Based on these findings and an assumption that values of the unsaturation index are not significantly altered in the sedimentary process, it has been proposed that stratigraphic profiles of U_{37}^k values down the core provide a temporal record of how sea surface temperature (SST) has changed in overlying waters on a relative³ and perhaps an absolute⁴ basis. Long-chain alkenones in the

f-turbidite from MAP provide strong evidence to validate the assumption that values of U_{37}^k are not significantly altered in the sedimentary process, thus supporting the proposed applications^{3,4}. Despite a major loss (~88%) in total alkenone content as a consequence of oxidative degradation, measured values of U_{37}^k in the oxidized and unoxidized intervals of the f-turbidite from cores 86P5 and 86P25 are practically identical (Table 1).

In calibrating the response of U_{37}^k values to growth temperature, Prah1 *et al.*⁵ monitored two additional compositional properties of the long-chain lipid series biosynthesized by *E. huxleyi*. These properties were the carbon chain length (K37/K38) in the C₃₇, C₃₈ and C₃₉ alkenone series and the distribution of di-unsaturated C₃₆ fatty acid ethyl esters (alkenoates) to alkenones (EE/K37). K37/K38 and EE/K37 both vary non-linearly with growth temperature (Fig. 1a and b, open symbols). Values of U_{37}^k , K37/K38 and EE/K37 were measured in MAP samples (Table 1) and in a variety of other sediment samples ranging in age from recent to late Pleistocene (≤ 26 kyr) and collected beneath warm tropical (MANOP Site C) to cold temperate (Multitracers and Washington coastal) waters. Values of K37/K38 and EE/K37 for these sediment samples (Fig. 1a and b, respectively, solid symbols) plotted against growth temperatures predicted from U_{37}^k values both show a trend paralleling the temperature calibration established for the laboratory culture. However, the sediment data are significantly offset from the laboratory calibration curves. Prah1

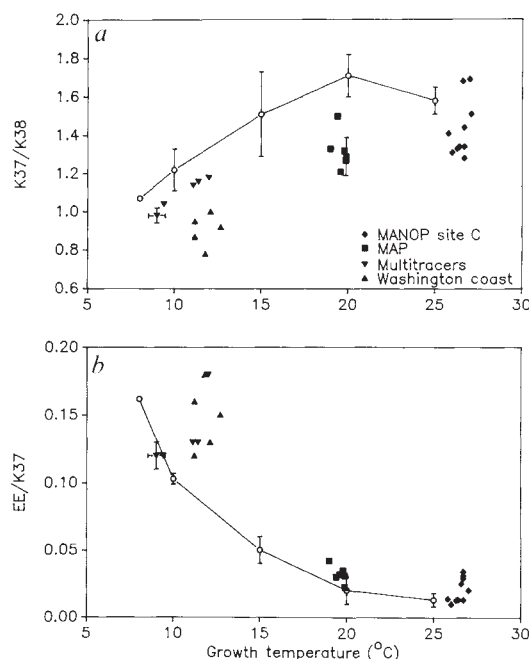


FIG. 1 Plots of a, the carbon-chain-length distribution (K37/K38) of the long-chain alkenone series and b, the di-unsaturated C₃₆ fatty acid ethyl ester (alkenoate) to alkenone composition (EE/K37) measured in laboratory cultures of *E. huxleyi* grown at various water temperatures (open symbols; data from ref. 5). Compositional data for the long-chain alkenone/alkenoate series measured in sediments from the continental shelf and slope off Washington state (Washington coast: 46° to 47° N, 124° 26' to 125° 5' W; 73–2,235 m water depth), in sediments along an east-west transect (~42° N) extending from 126° to 132° W off northern California (Multitracers: 2,500–3,000 m water depth), in sediments from the northern equatorial Pacific (MANOP site C: 1° N, 140° W; 4,400-m water depth) and in sediments from the f-turbidite of the Madeira Abyssal Plain (MAP) are also plotted (solid symbols). Values of K37/K38 and EE/K37 measured in sediments from these different marine environments are plotted at growth temperatures predicted from values of the unsaturation index U_{37}^k measured in each mixture using the following calibration equation: $U_{37}^k = 0.034 \times T + 0.039$ (ref. 5).

*et al.*⁵ noted this offset between laboratory and field data before the analyses of sediment samples from MAP and the east-west transect off northern California (Fig. 1). The offset could reflect a simple physiological difference between the laboratory strain of *E. huxleyi* used to calibrate these compositional properties and the average prymnesiophyte species contributing long-chain alkenones and alkenoates to marine sediments. Alternatively, it could reflect selective degradation of the tri-unsaturated compounds in the alkenone series. Here we show that $U_{37}^{k'}$ values are virtually unchanged despite major oxidative degradation of the total alkenone signal in the f-turbidite (Table 1). Thus, the second explanation for the observed offset (Fig. 1a and b) is excluded. Furthermore, because the values for K37/K38 and EE/K37 in MAP sediments from the f-turbidite show little difference between oxidized and unoxidized intervals (Table 1), the discrepancy between the laboratory calibration curves and field measurements would now seem attributable to differences in the field and laboratory phytoplankton specimens.

The insensitivity of the unsaturation index to change despite extensive oxidative degradation of the overall alkenone signal in the surface of the f-turbidite is an ideal feature if values of $U_{37}^{k'}$ are to be useful palaeotemperature indicators. Further understanding of the ecology of prymnesiophyte species contributing to the alkenone record in marine sediments is needed, however, to aid an accurate palaeoceanographic interpretation of stratigraphic records of $U_{37}^{k'}$ with depth in deep-sea sediment cores. Because the source of these biomarkers is limited to particular photosynthetic algae², production of these compounds in the ocean is confined to the euphotic zone. Nonetheless, this zone is not necessarily isothermal and may include the top portion of the thermocline. *E. huxleyi* contains the antenna pigment, 19'-hexanoyloxyfucoxanthin, which allows it to photosynthesize at wavelengths between 500–550 nm (ref. 22). Thus, these algae are not restricted to the surface layers of the euphotic zone, but may be adapted to deeper portions of this layer where light intensities are lower, light quality is shifted toward shorter wavelengths and water temperatures are colder than at the sea surface. In addition, we have not established the extent to which water temperatures predicted from $U_{37}^{k'}$ values are biased by the seasonality of the source input to sediments. Accurate palaeoclimatic interpretation of SST curves constructed from the down-core sediment record for $U_{37}^{k'}$ values will necessitate a detailed knowledge of what upper-ocean water temperature is recorded by the internal composition of this biomarker series. □

Benthic–pelagic coupling in a deep-sea benthic community

Gerhard Graf

Institut für Meereskunde, Düsternbrooker Weg 20, 2300 Kiel, FRG

SEASONAL productivity in shallow waters elicits a seasonal pattern of activity in the benthic community¹. This implies a very rapid response to fluctuations in the food supply to the benthos. Although there is a seasonal component to oxygen consumption on the deep-sea floor² and a rapid response of epibenthic foraminifera³ and other microorganisms to influx of phytodetritus⁴, the timescale of the response of organisms that inhabit the sediment itself has not been established. Here I present evidence that the response by a deep-sea benthic community to a pulse of natural organic matter occurs within days. In terms of activity and biomass, this rapid response was evident to a sediment depth of 9 cm.

The investigation was carried out on the Vøring plateau off the Norwegian continental margin at a water depth of 1,430 m (67°39'1" N, 05°48'3" E). The station was close to an escarpment at the foot of a 200 m-high ridge, where a multiple-sample sediment trap was moored during the experimental period⁵. Box corer samples were taken during the spring of 1986 and one during the summer of 1985.

From 24 May to 4 June, there was a sudden increase in chlorophyll *a* (ref. 5; Fig. 1) within the sediment itself, to a depth of 9 cm. During the rest of June there was a decrease in the concentration of chlorophyll *a*. This pattern was created by a pulse of faecal pellets which, especially at the end of May, were mainly derived from the copepod *Calanus finmarchicus*. Pellets accounted for 92% of the total carbon in the settling particles⁶. The ratio of particulate organic carbon (POC) to chlorophyll *a*-equivalent = 161 and the ratio of POC to particulate organic nitrogen (PON) = 7.2. These ratios were indicative of high quality food because they are close to those of fresh phytodetritus. During July there was an increase in the particulate influx of POC from 3 to 45 mg carbon m⁻² d⁻¹ (ref. 7). The same pattern was observed ~100 miles further north in the Lofoten Basin⁸. During July, however, there was a decrease in food quality, with POC: chlorophyll *a*-equivalent > 3,000, and POC: PON > 10. Consequently, the much higher POC flux caused only a minor increase of chlorophyll *a* within the upper layer of the sediment. The chlorophyll *a* pattern in the sediment (Fig. 1) thus reflected the same sedimentation pattern as that found in the trap. The integrated pellet pulse produced an increase in the concentration of chlorophyll *a* to 3.3 µg cm⁻², which was 82 times more than that collected in the trap. This agreed, however, with the estimated loss of 3.0 µg chlorophyll *a* per cm⁻² from the water column during the same period⁷. The discrepancy was caused by the poison in the trap, in this case chloroform, which extracted pigments from the sampled material and hence led to a gross underestimate of the input of chlorophyll *a*.

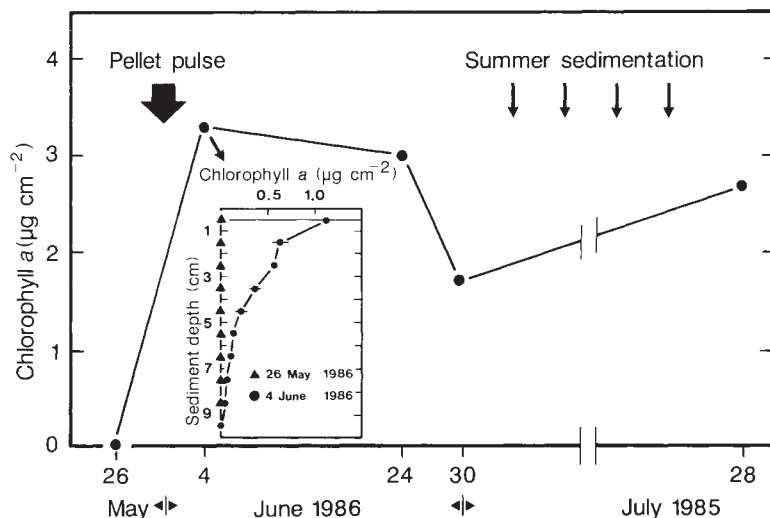
The chlorophyll *a* profile on 4 June implies that the pellets were incorporated into the surface sediment at a very rapid rate (> 1 cm d⁻¹), which was much more than could be expected for deep-sea bioturbation. The reason for this is that particle transport was not an effect of reworking, as recently described for echinurids⁹, but a unidirectional transport of food particles down a burrow by the sipunculans *Nephasoma* sp.. Therefore, it was inappropriate to apply Fick's law and calculate an apparent diffusion coefficient. These organisms build up to 11,000 burrows per m² in this area¹⁰. With their 6-mm-long introvert grazing on the sediment, the sipunculans can reach a 100% coverage of the surface. At a sediment depth of 9 cm ~30 burrows of *Nephasoma* sp. merge; from here one burrow leads off to a depth of more than 50 cm (ref. 10). Thus specimens were able

Received 15 May; accepted 11 August 1989.

- Okada, H. & Honjo, S. *Deep Sea Res.* **26**, 355–374 (1973).
- Marlowe, I. T. *et al. Br. phycol. J.* **19**, 203–216 (1984).
- Brassell, S. C. *et al. Nature* **320**, 129–133 (1986).
- Prahl, F. G. & Wakeham, S. G. *Nature* **330**, 367–369 (1987).
- Prahl, F. G., Muehlhausen, L. A. & Zahnle, D. L. *Geochim. cosmochim. Acta* **52**, 2303–2310 (1988).
- Farrington, J. W. *et al. Org. Geochem.* **13**, 607–617 (1988).
- Jasper, J. P. thesis, Woods Hole Oceanographic Institution (1988).
- Prahl, F. G., Muehlhausen, L. A. & Lyle, M. W. *Paleoceanography* (in the press).
- Wilson, T. R. S. *et al. Geochim. cosmochim. Acta* **49**, 811–822 (1985).
- Buckley, D. E. & Cranston, R. E. *Geochim. cosmochim. Acta* **52**, 2925–2939 (1988).
- de Lange, G. J., Jarvis, I. & Kuipers, A. in *Geology and Geochemistry of Abyssal Plains*, *Geol. Soc., Spec. publ.* **31**, 147–161 (1987).
- Weaver, P. E. & Kuipers, A. *Nature* **306**, 360–363 (1983).
- Weaver, P. E. & Rothwell, R. G. in *Geology and Geochemistry of Abyssal Plains*, *Geol. Soc., Spec. publ.* **31**, 71–86 (1987).
- McCave, I. N. & Jones, K. P. N. *Nature* **333**, 250–252 (1988).
- Thomson, J. *et al. in Geology and Geochemistry of Abyssal Plains*, *Geol. Soc., Spec. publ.* **31**, 167–177 (1987).
- ten Haven, H. L. *et al. Mar. Geol.* **75**, 137–156 (1987).
- Volkman, J. K. *et al. in Advances in Organic Geochemistry 1981* (ed. Bjoroy, M.) 228–240 (Wiley, 1983).
- Prahl, F. G. & Muehlhausen, L. A. in *Productivity in the Oceans: Present and Past* (eds Berger, W. H., Smetacek, V. S. & Wefer, G.) 271–289 (Wiley, 1989).
- Suess, E. *Nature* **288**, 260–263 (1980).
- Emerson, S. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaen to Present* (eds Sundquist, E. T. & Broecker, W. S.) 78–87 (American Geophysical Union, 1985).
- Finney, B. P., Lyle, M. W. & Heath, G. R. *Paleoceanography* **3**, 169–189 (1988).
- Haxo, F. T. *J. Phycol.* **21**, 282–287 (1985).
- Weliky, K. *et al. Limnol. Oceanogr.* **28**, 1252–1259 (1983).

ACKNOWLEDGEMENTS. We thank the NSF (Marine Chemistry and Marine Geology and Geophysics Divisions) and the Director of the Dutch Geological Survey for supporting this research.

FIG. 1 Concentrations of chlorophyll *a* in the sediment during spring 1986, collected with a box corer. The depth profiles indicate the fast penetration of pigments into the sediment. The integrated concentrations of chlorophyll *a* depict sedimentation of a pellet pulse deriving from *Calanus finmarchicus* grazing on the spring bloom.



to transport pelagic material, which settled on the sediment surface like foraminiferal shells or coccolithophores, quickly into the underlying geological strata.

The concentration of ATP in the sediment also responded to the pellet pulse (Fig. 2). At the end of May, concentrations of ATP were low and comparable to those of the North Pacific¹¹. A remarkable subsurface maximum was found, however, which corresponded to the depth range of the sipunculans¹⁰. After the input of food, an increase in ATP was visible to a depth of 10 cm, but was also especially pronounced at the sediment surface, where the elevated level persisted until 24 June. In two of the four cores a second ATP maximum was found between a depth of 6 and 8 cm. It corresponded to a horizontal burrow system of an enteropneust, *Stereobalanus canadensis*¹², which creates a special microhabitat where the wall lining is intensively explored by small agglutinated foraminifera (Jensen, personal communication). An ATP maximum was not found in every core because the subsurface burrow system covered at most 50 % of the area.

In shallow waters, changes in the concentrations of ATP may be used to calculate benthic production. Off the Vøring plateau the ATP increase may have also been derived from (1) microorganisms attached to the pellets, and (2) new ATP produced by organisms that had activated their metabolism after periods of starvation or dormancy. In the latter case there was no constant ratio between biomass and the concentration of ATP.

Shipboard measurements of oxygen consumption (Fig. 3) agreed with the depth-dependent rates measured *in situ* for the north-west Atlantic¹³, indicating that pressure effects of bringing the samples from the ambient water depth were not severe. No

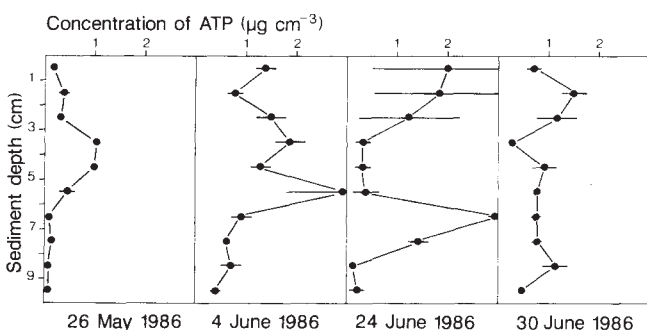


FIG. 2 The concentrations of ATP in the sediment ($n=5$) increased after the pellet pulse, especially in the surface layer. Two subsurface maxima correspond in sequence to the occurrence of sipunculids and enteropneusts, respectively.

immediate response to the pellet pulse was detected in early June, indicating that either a shorter event occurring later in May had been missed — an explanation suggested by laboratory simulation of sedimentation events¹⁴ — or the response was slow and did not occur before 30 June. Such a late response, however, would be indistinguishable from the effect of an increase in summer sedimentation.

We measured heat production by direct microcalorimetry¹⁵ at -0.5°C , close to the *in situ* temperature of -0.7°C (Fig. 4). Heat production in sediments is of mainly biological origin¹⁶ and is a direct measure of energy flow¹⁴. The heat production profile from 24 June shows that biological activity in the sediment was by no means restricted to the upper few millimetres of the sediment, as suggested by oxygen microprofile measurements in other deep sea areas¹⁷, but penetrated substantially below. If decomposition of organic matter is spread over a sediment depth of 9 cm — as was the case here — and is not supplied by a steady flow of particles but by seasonal pulses, a significant time lag between biological effects in the sediment and oxygen consumption at the surface is to be expected.

The response time of the benthic community was less than eight days. Future investigations of benthic-pelagic coupling in the deep sea have to be carried out on short timescales and include the whole sediment layer influenced by macrofauna. The fast deep-reaching response indicates that nutrient profiles should also be influenced. Another consequence of these results is that because pelagically produced particles arriving at the sea floor in association with a food pulse are deeply mixed with older material, they have little chance of becoming recorded in the sediment surface. □

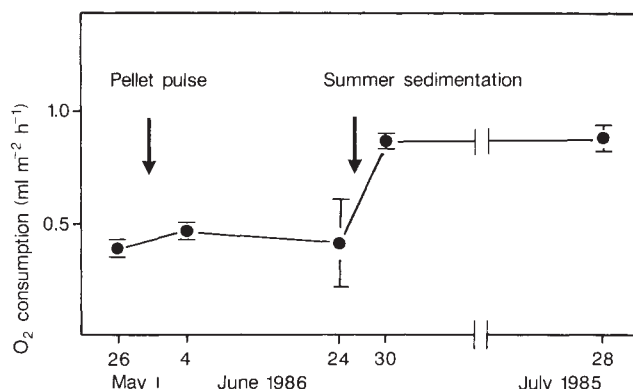


FIG. 3 Oxygen consumption of the sediment ($n=3$) did not show an immediate response to the food supply.

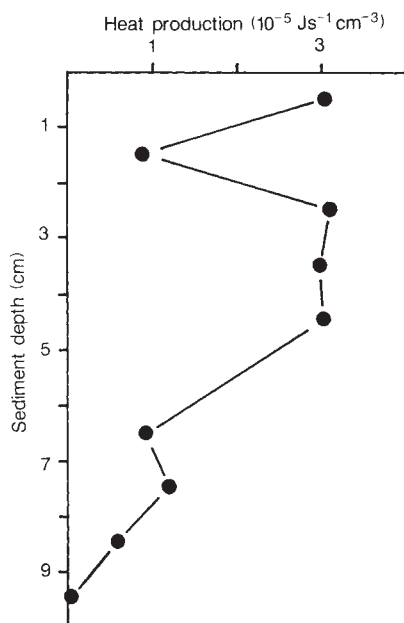


FIG. 4 Heat production in the sediment on 24 June; the decomposition of organic matter mainly occurs in the surface layer and in the subsurface layers influenced by the sipunculids.

Received 22 May; accepted 31 July 1989.

- Graf, G., Schultz, R., Peinert, R. & Meyer-Reil, L. -A. *Mar. Biol.* **77**, 235-246 (1983).
- Smith, K. L. *Limnol. Oceanogr.* **32**, 201-220 (1987).
- Gooday, A. J. *Nature* **332**, 70-73 (1988).
- Lochte, K. & Turley, C. M. *Nature* **333**, 67-69 (1988).
- Jeffrey, S. W. & Humphrey, G. F. *Biochem. Physiol. Pflanzen* **167** (1975).
- Bathmann, U. V., Noji, T. T., Voss, M., & Peinert, R. *Mar. Ecol. Prog. Ser.* **38**, 45-51 (1987).
- Peinert, R., Bathmann, U., Bodungen, B. v. & Noji, T. *Mitt. geol. paläont. Inst. Univ. Hamb.* **56**, 149-164 (1987).
- Wefer, G. *Terra cognita* **5**, 90 (1985).
- Smith, C. R., Jumars, P. A. & DeMaster, D. J. *Nature* **323**, 251-253 (1986).
- Romero-Wetzel, M. B. *Mar. Biol.* **96**, 87-91 (1987).
- Craven, D. B., Jahnke, R. A. & Carlucci, A. F. *Deep Sea Res.* **33**, 379-390 (1986).
- Romero-Wetzel, M. B. *Sarsia* **74**, 85-89 (1989).
- Smith, K. L. *Mar. Biol.* **47**, 337-347 (1978).
- Graf, G. *Mar. Ecol. Prog. Ser.* **39**, 23-29 (1987).
- Graf, G., Martens, V., Queisser, W., Weinholz, P. & Altenbach, A. *Mar. Ecol. Prog. Ser.* **45**, 201-204 (1988).
- Pamatmat, M. M. *Science* **215**, 395-397 (1982).
- Reimers, C. E., Fischer, K. M., Merewether, R. Smith, K. L. & Jahnke, R. A. *Nature* **320**, 741-744 (1986).

ACKNOWLEDGEMENTS. I thank A. Scheitz for technical assistance and Drs U. Bathmann, R. Lampitt and E. Suess for comments on the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft.

A direct nicotinic receptor-mediated inhibition recorded intracellularly *in vitro*

Linda A. Wong & Joel P. Gallagher

Department of Pharmacology and Toxicology, University of Texas Medical Branch, Galveston, Texas 77550, USA

ACETYLCHOLINE activates both nicotinic and muscarinic receptors in the central nervous system¹. Although the action of acetylcholine at muscarinic receptors has been well characterized, relatively little is known at the cellular level concerning nicotinic receptor stimulation in brain. Central nicotinic receptors have been implicated in Alzheimer's disease², seizure activity³, the generation of slow-wave theta rhythm in the hippocampus⁴ and the potential abuse liability of nicotine⁵. At the neuronal level, nicotinic agonists have been most often associated with postsynaptically mediated

excitation and membrane depolarization at various sites, including Renshaw spinal motoneurons⁶, locus coeruleus⁷ and the medial habenular nucleus⁸. Nicotine acting presynaptically can produce either excitation or inhibition indirectly through the release of endogenous transmitters or modulators⁹⁻¹². Whereas a direct inhibitory effect of nicotine has been suggested by one *in vivo* extracellular recording study in rat cerebellar Purkinje neurons¹³, the mechanism(s) underlying this action is not yet known. We now report our findings obtained using *in vitro* intracellular methods in a submerged brain slice preparation in which application of nicotinic agonists to rat dorsolateral septal neurons reveal a direct membrane hyperpolarization mediated by an increase in potassium conductance.

The dorsolateral septal nucleus (DLSN) lies within the septohippocampal circuitry and is believed to play a part in integrating afferent electrical activity arising from the hippocampus before terminating in the medial septum/nucleus of Diagonal Band of Broca complex (MS/DBB)¹⁴⁻¹⁶. The MS/DBB is, in turn, involved in the generation and maintenance of hippocampal theta rhythm¹⁷⁻¹⁹.

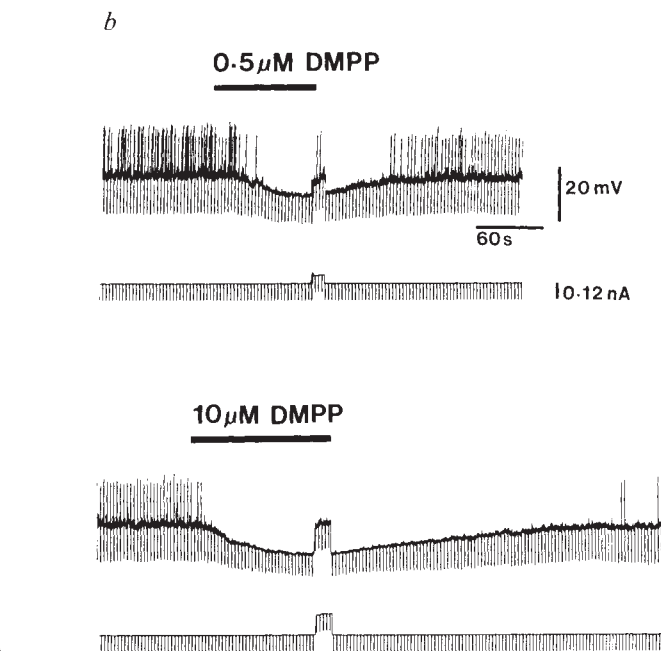
Several lines of evidence point to a role for afferent cholinergic pathways in the function of lateral septal nuclei. Acetylcholinesterase staining^{20,21} as well as nicotinic and muscarinic binding sites²²⁻²⁵ are evident in this region. Additionally, immunohistochemical studies indicate that the DLSN receives a cholinergic innervation from the laterodorsal tegmental²⁶ and pedunculopontine nuclei²⁷. Consistent with these anatomical findings, electrophoretic application of acetylcholine (ACh) *in vivo* depresses spontaneous neuronal activity and electrically evoked field potentials in lateral septum, suggesting that ACh may serve to inhibit synaptic transmission in this region^{28,29}. Also, recent studies from this laboratory, using *in vitro* intracellular recording methods, have provided direct evidence that neurons in the DLSN are cholinceptive^{30,31}. These latter findings indicate that cholinergic agonists produce effects which are mediated in part at pre- and postsynaptic M₁ muscarinic receptors in the DLSN. The responses to non-selective cholinergic agonists such as ACh or carbachol, however, were not completely antagonized by atropine, implying the presence of nicotinic receptors. In this study we investigated the membrane effects and ionic mechanism underlying nicotinic receptor activation in the DLSN using ACh (tested in the presence of 5 μ M atropine), nicotine, and 1,1-dimethyl-4-phenylpiperazinium (DMPP) as specific agonists.

Intracellular recordings were obtained from the DLSN neurons contained within a submerged brain slice, continuously superfused with Krebs buffer. Data were collected from 105 DLSN neurons which exhibited a resting membrane potential of -59.9 ± 0.4 mV (mean $\pm$ s.e.m.; $n = 55$), a membrane time constant of 29.9 ± 2.1 ms ($n = 30$), and an input resistance of 99.6 ± 4.9 M Ω ($n = 30$). The method of pressure ejection was frequently used in this study to apply nicotinic agonists to minimize receptor desensitization resulting from bath application of agonists. DMPP was used as a prototypic nicotinic agonist in our study as it was found to be more potent and less desensitizing than nicotine in our preparation (unpublished data). Similar observations regarding differences in the potency between DMPP and nicotine have been noted in studies of the peripheral ganglia³².

Application of DMPP (5-10 mM) by pressure-ejection produced a membrane hyperpolarization (6.0 ± 0.4 mV) associated with decreased membrane resistance and cessation of spontaneous firing in 43 out of 61 neurons (Fig. 1a, top). This membrane hyperpolarizing response could be mimicked by pressure application of ACh (100 mM) in four out of eight neurons (4.7 ± 1.2 mV; Fig. 1a, middle) or nicotine (100 mM) in eight out of sixteen neurons (6.0 ± 1.2 mV; Fig. 1a, bottom). In addition, a small amplitude depolarization (1-2 mV) occasionally followed the membrane hyperpolarization induced by nicotine (Fig. 1a, bottom). Superfusion of DMPP (0.1-100 μ M;

FIG. 1 Effects of DMPP, ACh and nicotine (Nic) on the membrane potential and input conductance of DLSN neurons. *a*, Pressure application of DMPP (10 mM, top); ACh (100 mM, in the presence of 5 μ M atropine to block muscarinic receptors, centre); and Nic (100 mM, bottom) produced membrane hyperpolarizations (4–6 mV) lasting 20–60 s. A small amplitude membrane depolarization follows the prominent Nic-induced membrane hyperpolarization. The neurons resting membrane potential was –56 mV (top), –59 mV (centre), and –58 mV (bottom), respectively. *b*, The effect of DMPP is concentration-dependent. Bath-applied DMPP (0.5 μ M) produced a 7 mV membrane hyperpolarization accompanied by decreased input resistance. Superfusion of 10 μ M DMPP, in the same cell, produced a larger amplitude hyperpolarization (12 mV). The decrease in membrane resistance and inhibition of firing persisted during return of membrane potential to control level from the peak of hyperpolarization by injection of a depolarizing DC current. Upward deflections in upper trace represent action potentials, whereas downward deflections represent voltage responses to constant current injection, shown in lower trace. Resting membrane potential, –60 mV.

METHODS. Rat septal brain slices were prepared as previously described³⁵. Male Sprague-Dawley rats weighing 150–250 g were decapitated, the brain rapidly dissected out and placed in cold Krebs solution of the following composition (mM): NaCl, 117; KCl, 4.7; MgSO_4 , 1.2; CaCl_2 , 2.5; NaH_2PO_4 , 1.2; glucose, 11.5; NaHCO_3 , 25; pre-bubbled with 95% O_2 and 5% CO_2 . Transverse blocks of tissue containing the septum were serially cut on a Vibroslice (Camden Instrument) into 500 μ m sections. A single slice was placed in the recording chamber and superfused with oxygenated Krebs solution warmed to $32 \pm 1^\circ\text{C}$. Conventional intracellular recording methods were employed using glass micropipettes (75–110 $\text{M}\Omega$) filled with 4 M potassium acetate or 2 M potassium chloride. Hyperpolarizing current pulses (0.1–



0.15 nA) were injected into the recording pipette by commands generated through an Axoclamp (Axon Instruments) amplifier. Orthodromic stimuli were delivered with square-wave pulses (5–20 V, 0.15 ms) through concentric bipolar electrodes placed focally within the DLSN. DMPP (5–10 mM), ACh chloride (100 mM) and nicotine hydrogen tartrate (100 mM) were applied by pressure-ejection (15–20 psi, 20–50 ms; Picospritzer, General Valve Corporation) through a glass micropipette (tip diameter 20–35 μ m) positioned in solution within 200–300 μ m from the recording electrode above the surface of the slice. By this method of application the agonist was diluted by a factor of about 1/100. All other drugs were freshly dissolved in Krebs buffer and applied by superfusion. Atropine (5 μ M) was included in the bathing medium to block muscarinic receptor-mediated effects during ACh-application. Only one neuron was obtained from each brain slice for pharmacological testing.

FIG. 2 Postsynaptic site of DMPP-induced membrane hyperpolarization. Slices were treated with TTX (1 μ M) or zero Ca^{2+} plus 6 mM Mg^{2+} , which completely blocked orthodromically stimulated synaptic responses after 1 and 4 min, respectively. Pressure-applied DMPP (10 mM) produced a membrane hyperpolarization that persisted in (a) TTX (1 μ M) or (b) zero Ca^{2+} plus 6 mM Mg^{2+} . The evoked synaptic potentials are represented by the additional set of downward deflections in the control voltage traces. In *b*, the bottom panel shows expanded traces (labelled 1 and 2) of the synaptic waveform evoked by orthodromic stimuli (20 V, 0.15 ms). Control synaptic waveforms (labelled 1) consist of a fast inhibitory potential (f -i.p.s.p.) mediated mainly by GABA_A receptors and a late hyperpolarizing potential (LHP), which is mediated in part by GABA acting on GABA_B receptors³⁵. Both f -i.p.s.p. and LHP were completely abolished in calcium-free medium (see label 2), leaving a stimulus artefact denoted by the short vertical deflection in the trace. Further increase of stimulus intensity to a supramaximal level (60 V) did not restore either the f -i.p.s.p. or LHP potentials, indicating a complete blockade of synaptic transmission in calcium-free medium.

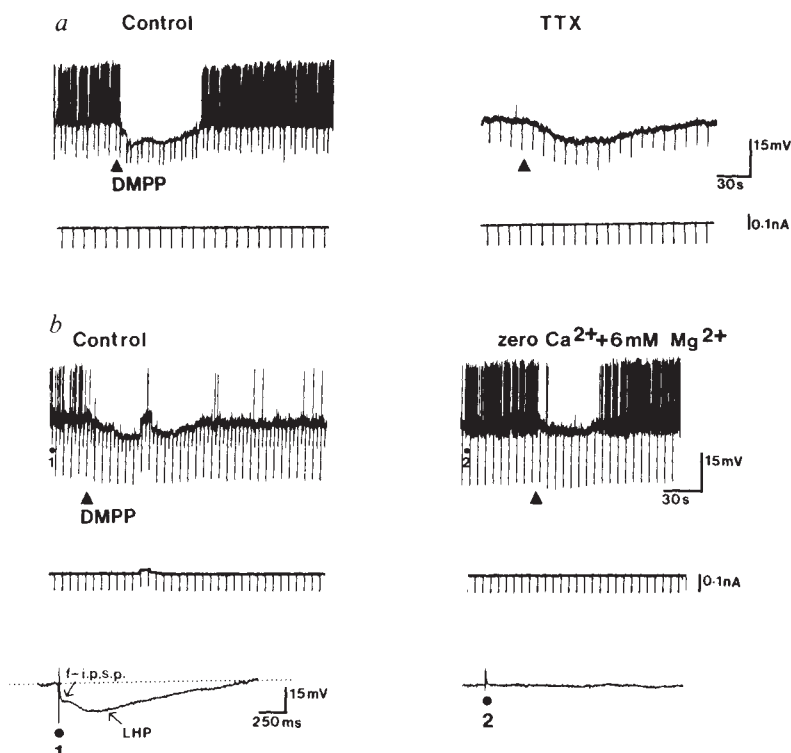


Fig. 1b) also produced membrane hyperpolarization (2–13 mV) in 17 of 20 cells tested. The amplitude and duration of the hyperpolarization depended on the concentration of DMPP applied.

To determine whether the nicotinic agonist-induced membrane hyperpolarization resulted from a direct action on the postsynaptic membrane, DMPP was applied in the presence of either tetrodotoxin (TTX; 0.1–1 μ M) or a modified Krebs solution containing no added calcium and 6 mM magnesium (zero Ca^{2+} , 6 mM Mg^{2+}) which blocked synaptic transmission (Fig. 2). Pressure-applied DMPP produced a membrane hyperpolarizing response that persisted in TTX ($n=8$; Fig. 2a) or zero Ca^{2+} , 6 mM Mg^{2+} medium ($n=4$; Fig. 2b), indicating that DMPP was activating receptors located directly on the neuron studied. It is interesting to note that although the response to DMPP persisted in zero Ca^{2+} , 6 mM Mg^{2+} medium (Fig. 2b), at which time orthodromic synaptic responses were completely abolished (Fig. 2b, bottom panel), the amplitude of the DMPP-induced response appeared attenuated. Although the depression of the DMPP-mediated hyperpolarization may be due to an incomplete block of synaptically released neurotransmitter, it is more probable that the postsynaptic response to nicotinic agonists in DLSN is in some way dependent on Ca^{2+} or blocked by Mg^{2+} . Preliminary results using intracellular injection of the

calcium chelator, EGTA, reveal that the DMPP-induced membrane hyperpolarization is completely blocked (data not shown).

The membrane hyperpolarizing action of DMPP, ACh, and nicotine was accompanied by a decrease of input resistance indicating an increase of ionic conductance. This agonist-induced hyperpolarization was reproducible whether the electrode contained potassium chloride ($n=3$) or potassium acetate ($n=72$), suggesting that Cl^- ions were not involved. A plot of current-voltage relationships recorded before application of DMPP and at the peak of the DMPP-induced membrane hyperpolarization (Fig. 3a) revealed a reversal potential of -91.2 ± 1.9 mV ($n=8$), a value close to the equilibrium potential for K^+ ion. To obtain further evidence that the potential changes were due only to changes in K^+ conductance, DMPP was applied in the presence of altered K^+ ion concentration. The DMPP-induced membrane hyperpolarizing response reversed at -92 ± 1.5 mV ($n=3$) and -77 ± 1.2 mV ($n=3$) in 4.7 mM and 8.0 mM K^+ , respectively (Fig. 3b). In each case, recovery followed wash-out of 8 mM K^+ and the reversal potential after return to control 4.7 mM K^+ was -91 ± 1.4 ($n=3$). This shift in the reversal potential of the hyperpolarization was close to that predicted by substitution of the altered extracellular K^+ concentration in the Nernst equation and provides strong evidence for the role of K^+ ions in mediating the membrane hyperpolarization produced by DMPP.

Our results demonstrate that ACh, nicotine and DMPP have direct actions on DLSN neurons, as indicated by persistence of the nicotinic-mediated membrane hyperpolarization in TTX or zero calcium, high magnesium solutions. To our knowledge, this is the first demonstration of a direct nicotinic receptor-mediated membrane hyperpolarization and associated inhibition of neuronal activity recorded intracellularly from a mammalian neuron. In addition, we have performed pharmacological tests to characterize this nicotinic-mediated response in the DLSN neurons (data not shown). These findings indicate that the membrane hyperpolarization produced by pressure-applied DMPP or nicotine can be effectively blocked by prior treatment of the slice with superfused mecamylamine (50–100 μ M), a ganglionic nicotinic antagonist, and κ -bungarotoxin (0.7–1.0 μ M), a snake venom toxin that has been shown to block neuronal nicotinic receptors^{33,34}.

As neurons in the DLSN are known to project to the MS-DBB complex^{14–16} and the latter is involved in the generation and/or maintenance of hippocampal theta rhythm^{17–19}, drugs or transmitters that alter the neuronal activity of DLSN may also influence MS/DBB activity and correspondingly, theta rhythm. The presence of cholinergic innervation to the lateral septum from brain stem nuclei^{27,28} also raises the possibility that brain stem areas may be involved in the modulation of septohippocampal theta rhythm through cholinergic mechanisms in the DLSN. It is possible that inhibition of DLSN neurons through nicotinic receptor stimulation may be one of the mechanisms by which ascending control of theta activity is facilitated. □

Received 24 April; accepted 15 August 1989.

- Krnjevic, K. *Physiol. Rev.* **54**, 418–540 (1974).
- Whitehouse, P. J. et al. *Arch. Neurol.* **45**, 722–724 (1988).
- Miner, L. L., Marks, M. J. & Collins, A. L. *J. Pharmacol. exp. Ther.* **239**, 853–860 (1986).
- Bland, B. H. & Colom, L. V. *Brain Res.* **440**, 167–171 (1988).
- Byrne, G. *Science* **240**, 1143 (1988).
- Curtis, D. R. & Ryall, R. W. *Exp. Brain Res.* **2**, 6–80 (1966).
- Egan, T. M. & North, R. A. *Neuroscience* **19**, 565–571 (1986).
- McCormick, D. A. & Prince, D. A. *J. Neurosci.* **7**, 742–752 (1987).
- Chesselet, M. F. *Neuroscience* **12**, 347–375 (1984).
- Westfall, T. C. *Neuropharmacology* **13**, 693–700 (1974).
- Hery, F., Bourgoin, S., Hamon, M., Ternaux, J. P. & Glowinski, J. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **396**, 91–97 (1977).
- Freund, R. K., Jungschafer, D. A., Collins, A. C. & Wehner, J. M. *Brain Res.* **453**, 215–220 (1988).
- de la Garza, R., McGuire, T. J., Freedman, R. & Hoffer, B. J. *Neurosci. Lett.* **80**, 303–308 (1987).
- DeFrance, J. F. in *The Septal Nuclei* (ed. DeFrance, J. F.) 185–227 (Plenum, New York, 1976).
- McLennan, H. & Miller, J. J. *J. Physiol. (Lond.)* **237**, 607–624 (1974).
- Swanson, L. W. & Cowan, W. M. *J. comp. Neurol.* **186**, 621–656 (1979).
- Green, J. D. & Arduini, A. *J. Neurophysiol.* **17**, 553–557 (1954).
- Stumpf, C., Petsche, H. & Gogolak, G. *Electroenceph. clin. Neurophysiol.* **14**, 212–219 (1962).
- Bland, S. K. & Bland, B. H. *Brain Res.* **375**, 102–116 (1986).
- Harkmark, W., Melligren, S. I. & Srebro, B. *Brain Res.* **95**, 281–289 (1975).

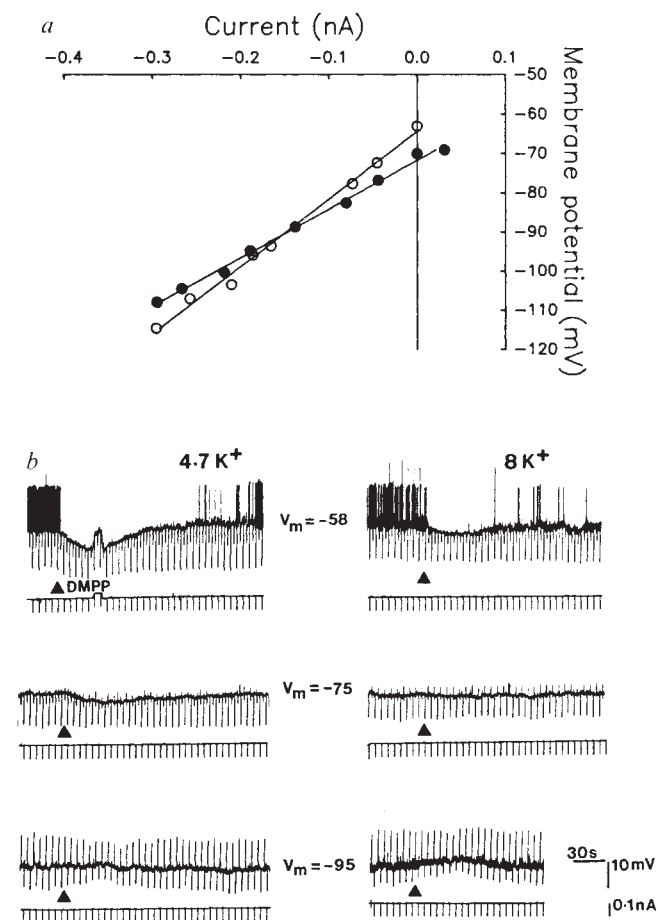


FIG. 3 Potassium dependence of DMPP-induced membrane hyperpolarization. *a*, A plot of current-voltage relationships obtained before application of DMPP (open circles) and at the peak of the response to superfused DMPP (10 μ M; closed circles) revealed a reversal potential of -90.5 mV. Lines were calculated using linear regression. *b*, DMPP was pressure-applied at holding potentials of either -58 , -75 or -95 mV in 4.7 mM and 8 mM external K^+ solutions. In 4.7 mM external K^+ , the DMPP-induced membrane hyperpolarization exhibited an apparent reversal at -95 mV which was shifted to -75 mV in 8 mM external K^+ . Evoked responses are represented by the additional sets of downward and upward deflections.

21. Paxinos, G. & Watson, C. *The Rat Brain in Stereotaxic Co-ordinates* Plate 17 (Academic, City, 1986).
22. Rainbow, T. C., Schwartz, R. D., Parsons, B. & Kellar, K. *Neurosci. Lett.* **50**, 193-196 (1984).
23. Clarke, P. B. S., Schwartz, R. D., Paul, S. M., Pert, C. B. & Pert, A. *J. Neurosci.* **5**, 1307-1315 (1985).
24. Schwartz, R. D. *Life Sci.* **38**, 2111-2119 (1986).
25. Nonaka, K. & Moroh, I. *Brain Res.* **296**, 295-303 (1984).
26. Satoh, K. & Fibiger, H. C. *J. comp. Neurol.* **253**, 277-302 (1986).
27. Hallanger, A. E. & Wainer, B. H. *J. comp. Neurol.* **274**, 483-515 (1988).
28. DeFrance, J. F., Yoshihara, H., McCrea, R. A. & Kitai, S. T. *Expl Neurol.* **48**, 502-523 (1975).
29. Segal, M. *Life Sci.* **14**, 1345-1351 (1974).
30. Hasuo, H., Gallagher, J. P. & Shinnick-Gallagher, P. *Brain Res.* **438**, 323-327 (1988).
31. Wong, L. A., Hasuo, H. & Gallagher, J. P. *Soc. Neurosci. Abstr.* **13**, 268 (1987).
32. Taylor, P. in *The Pharmacological Basis of Therapeutics* p 218 (Macmillan, New York, 1985).
33. Chiappinelli, V. A. *Pharmac. Ther.* **31**, 1-32 (1985).
34. Loring, R. H., Andrews, D., Lane, W. & Zigmond, R. E. *Brain Res.* **385**, 30-37 (1986).
35. Stevens, D. R., Gallagher, J. P. & Shinnick-Gallagher, P. *Synapse* **1**, 184-190 (1987).

ACKNOWLEDGEMENTS. We thank Drs Michael J. Twery and Kevin D. Phelan for review of this manuscript. This research was supported in part by the US Army Medical Research Development Command, the National Institutes of Mental Health and The Council for Tobacco Research.

Mutant *Drosophila* embryos in which all cells adopt a neural fate

Marc Bourouis, Pascal Heitzler, Mariame El Messal & Pat Simpson*

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Faculté de Médecine, 11 Rue Humann, 67085 Strasbourg Cédex, France

In the *Drosophila* embryo, early developmental decisions lead to all cells adopting one of several initial fates, such as those characteristic of the germ layers. The central nervous system is formed subsequently from the neurogenic region of the ectoderm, in which progenitor cells of the neuroblasts and ventral epidermis are

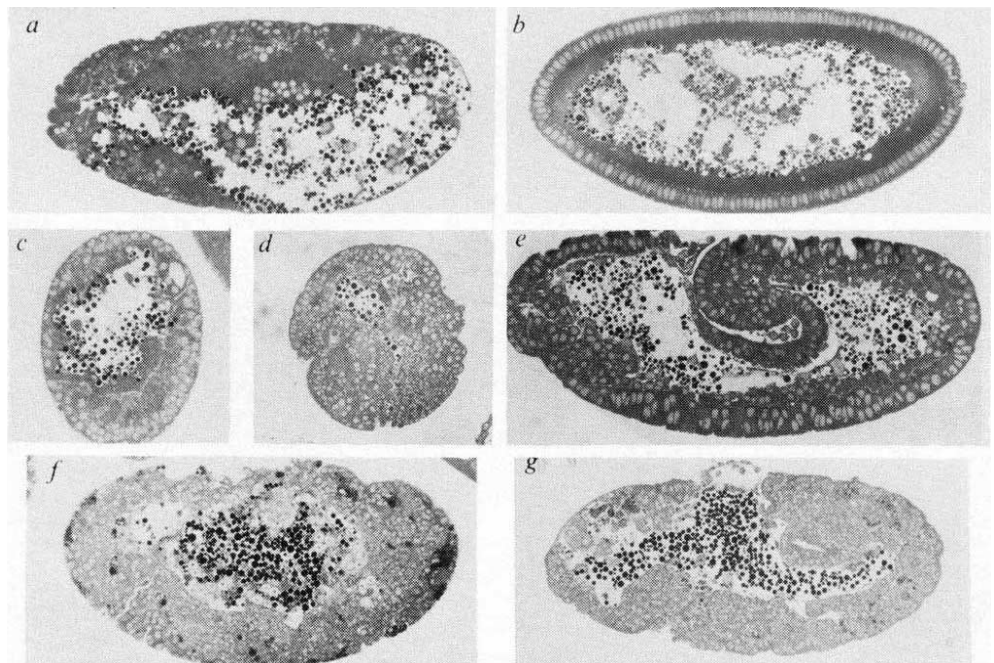
intermingled^{1,2}. Two classes of genes govern the segregation of neuroblasts and peripheral sensory organs. The pro-neural³ class of genes, for example, the achaete-scute complex, participates in the initial decision to make each uniquely positioned neuroblast or sensory organ⁴⁻⁷, but are initially expressed in groups of cells^{8,9}. The segregation of a neuroblast or sensory organ from an equivalent group¹⁰ of equipotential cells involves a mechanism of lateral inhibition whereby the future epidermal cells are prevented from engaging in the primary dominant neural fate^{2,11-15}. In the absence of this inhibitory signal, all cells of the group will become neural by default^{1,16-19}. The neurogenic class of genes is thought to mediate these cell interactions²⁰. Here we report that cells in embryos mutant for *shaggy* which are unable to adopt any of the early initial fates, instead develop neural characteristics.

Shaggy (*sgg*) null embryos produced by a loss-of-function mutation in both the female germ line and the zygote (see Fig. 1) were abnormal at a stage as early as syncytial blastoderm, when no clear cortical cytoplasm is evident in living embryos. Histological sections (Fig. 1) showed that nuclei migrate to the periphery but that cellularization is delayed. At the blastoderm stage the embryos did not form a continuous sheet of elongated cells. Instead cellularization was disordered, leading in some cases to embryos with gaps in the blastoderm. Furthermore, the cells were abnormally small and round. None of the normal events of gastrulation occurred and the embryos showed no sign of the development of an ectoderm, mesoderm or endoderm. Cells started to divide and their progeny accumulated in the interior of the egg. Over several hours the disorganized embryos became filled with small round uniform cells which resembled the ganglion mother cells that normally arise in the central nervous system. The neural character of the cells was confirmed using antibodies specific for neuronal cells. As shown in Fig. 2, most, if not all, of the cells stained positively with both an antibody against a neuronal cell-surface epitope, anti-horse-

* To whom correspondence should be addressed.

FIG. 1 Histological sections of mutant *sgg* embryos. *a*, A *sgg* null embryo showing the disordered cellularization observed at blastoderm. Note the abnormal segregation of cortical cytoplasm and yolk, the round peripheral cells and the presence of nuclei not incorporated into cells. *b*, Paternally rescued embryo showing an almost normal cellular blastoderm. Nuclei lagging behind those at the periphery are frequently encountered in these embryos. *c*, *d*, Transverse sections of *sgg* null embryos. The embryo in *c* was ~4-h-old, a stage at which normal embryos have gastrulated. Note that the cells around the entire periphery are internalizing. The embryo in *d* is slightly older and smaller cells are accumulating interiorly. *e*, Paternally rescued embryo at the end of the rapid phase of germ-band extension. These embryos undergo normal gastrulation and the three germ layers are visible. This implies that eggs made in the absence of the maternal *sgg* product have normal polarity, that is, the fate map is not changed. *f*, A *sgg* null embryo at a later stage, about the same age as the embryo in *g*. A study of serial sections shows that these embryos are like bags of marbles, being filled with round cells, interrupted only by clumps of unconsumed yolk. *g*, Paternally rescued embryo at the beginning of germ-band retraction. Hyperplasia of the developing central nervous system can be seen.

METHODS. Germ-line clones mutant for *sgg* (3B1) were generated by X-ray induced mitotic recombination. Irradiated *sgg/ovo*^{D1} (ref. 26) females were crossed with wild-type males to give progeny, half of which receives a paternally derived *sgg*⁺ allele. Observations of living and sectioned embryos shows a bimodal distribution. Embryos were fixed with heptane and glutaral-



dehyde and processed as previously described²⁷. Embryos were embedded in durcupan (Fluka, Switzerland) and semi-thin sections of 1.5-2.0 µm were stained with methylene blue. The allele *sgg*^{D127} was used for the histology but observations of living embryos and of the cuticular phenotypes of paternally rescued embryos for the alleles *sgg*^{D127}, *sgg*^{k22} and *sgg*^{b12} show identical phenotypes. Genetic analysis and preliminary molecular characterization (Yves Grau and P.S., unpublished results) indicates that both *sgg*^{D127} and *sgg*^{b12} are amorphic alleles.

radish peroxidase antibody (anti-HRP)²¹, and a very early neuronal nuclear marker, monoclonal antibody mAb44C11²², which stains the differentiating neurons in the wild-type embryo.

The defective cellularization observed in *sgg* null embryos implies that the *sgg* gene has an early activity, perhaps during cellularization itself. Note, however, that there is no absolute requirement for the *sgg* gene product for vital cellular functions; mutant cells are viable in the germ line, larva or imaginal discs,

and can proliferate normally²³.

We partially rescued the null phenotype of *sgg*-mutant embryos, selectively allowing either maternal or zygotic gene expression. This revealed the effects of *sgg* mutants on other developmental processes. The zygotically rescued embryos formed an almost normal blastoderm and gastrulate (Fig. 1). The embryos first deviated conspicuously from wild-type embryos at neurogenesis when a greater number of neuroblasts segregate.

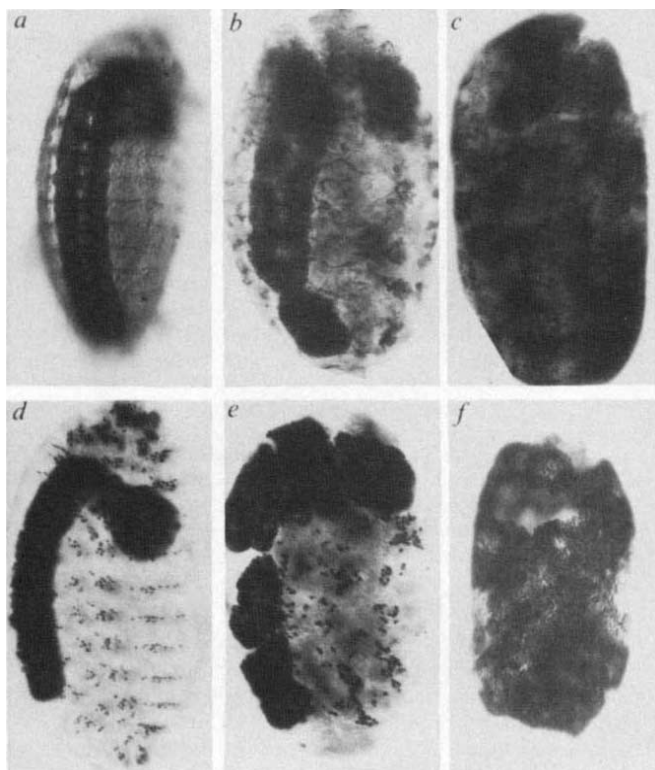
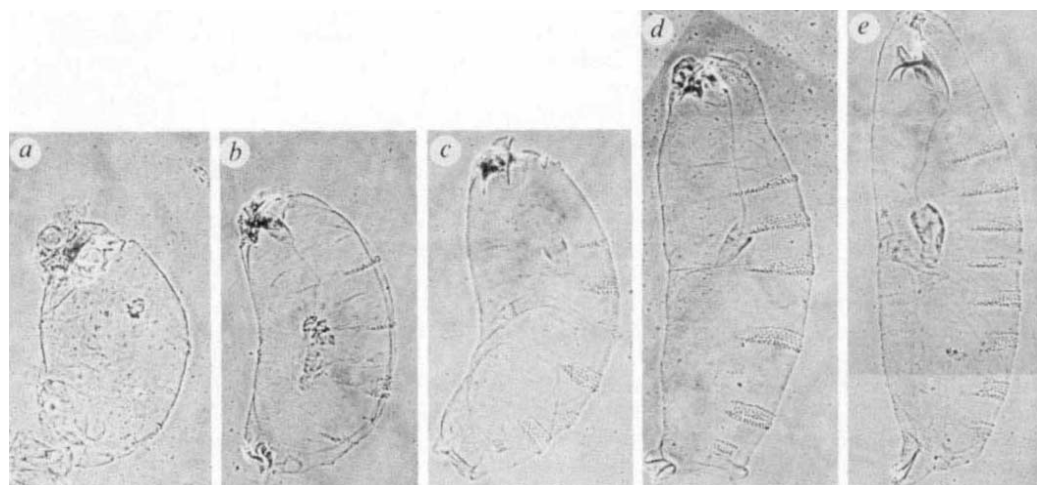


FIG. 2 Neural-specific immunocytochemical staining of *sgg* embryos selected at a stage when development of the central nervous system is complete. Left, wild-type embryos; middle, paternally rescued *sgg* embryos; right, *sgg* null embryos. *a, b, c*, Embryos stained with anti-HRP antibody²¹ (ventro-lateral view). *d, e, f*, Embryos stained with mAb44C11²² (lateral view). In the null embryos most, if not all, cells stain positively for both markers. In the paternally rescued embryos staining of both the central nervous system and the peripheral nervous system is evident. Enlargement of the ventral nerve cord reveals hyperplasia and also in *b*, local hyperplasia of the scolopidia of the chordotonal organs at the periphery.

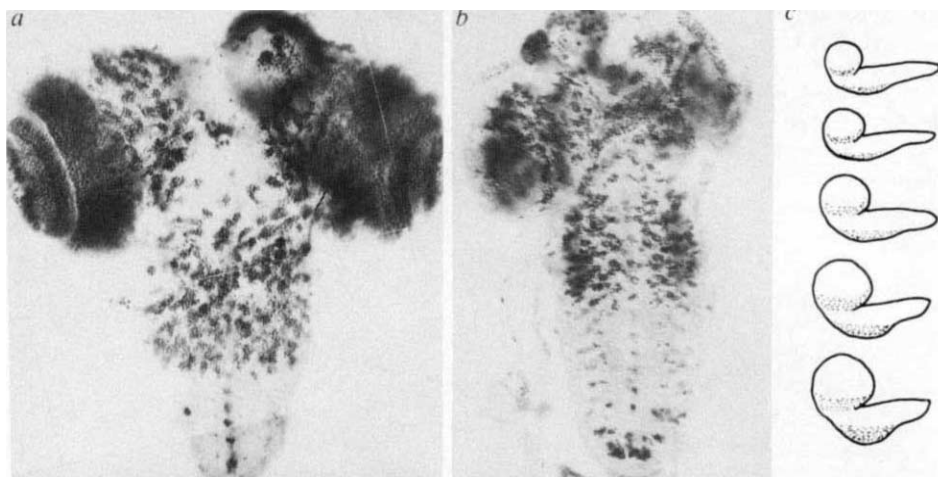
METHODS. Mutant embryos, obtained as described in Fig. 1, were treated as described previously⁷, using anti-HRP antibody (Sigma) and the mAb44C11. For mAb44C11 a biotinylated goat anti-mouse IgM (Amersham) together with the ABC kit (Vector Labs) was used for detection.

FIG. 3 Cuticular preparations of zygotically rescued *sgg*-mutant embryos. Females with a mutant germ line (see Fig. 1) were crossed either with $+Y$ males (*a*) or $+w^+Y$, *sgg*⁺; *dup(1; 2)w⁺70h*, *sgg*⁺/+ males (*b, c, d* and *e*). Paternally rescued embryos from the first cross secreted a very short cuticle that was organized in the form of a tube (*a*). Embryos of this genotype display neural hyperplasia (Figs. 1 and 2). Complete hyperplasia found locally would be expected to cause holes in the ventral cuticle, as in the neurogenic mutants^{16,17}. The presence of a continuous cuticle, therefore, indicated that the remaining epidermal cells were distributed along the entire length of the embryo. This implies an even distribution of the multitudinous neuroblasts, a situation reminiscent of the even spacing of bristles seen in the mutant imaginal epidermis¹⁵. The cuticle displayed a lawn of dorsal hairs, but ventrally most of the denticle belts were lacking. Superficially this phenotype resembled that of embryos mutant for *naked*²⁸, a gene of the segment polarity class. From the second cross zygotes may receive one or two paternal copies of *sgg*⁺, and in male embryos these will be dosage compensated. These embryos show more effective rescue (*b-e*). The embryo in *e* had an almost normal cuticular pattern. Progressively more bands of



ventral denticles appeared, beginning at the lateral margin. The pattern of rescue showed a marked regional preference and occurred in an alternating segment mode, the odd-numbered segments appearing before the even-numbered ones. Zygotic loss of *sgg* did not lead to abnormal segmentation of the cuticle (see Fig. 4). Therefore, the defects observed, which were caused by maternal loss, imply that the segmentation process requires a threshold level of a probably evenly distributed maternal product. Further investigation is necessary to determine whether the anomalous segmentation reflects a requirement for *sgg* specific to this patterning process or a more general requirement in the cells of the epidermis.

FIG. 4 Growth of larval nervous systems in maternally rescued *sgg* mutant zygotes. When derived from females with a wild-type germ line, mutant animals hatched as larvae. They displayed abnormal growth of the nervous system. *a*, *b*, Larval central nervous systems of late third-instar wild-type and mutant siblings labelled by incorporation of 5-bromodeoxyuridine to reveal neuroblasts and their progeny (dorsal view). Such dividing cells produce the future adult central nervous system. The cell clusters in these figures represent the progeny of individual neuroblasts²⁹. Cell counts from a number of such preparations and a parallel study of staged toluidine-blue-stained preparations²⁹ (data not shown) showed that similar numbers of neuroblasts were present in mutant and wild-type central nervous systems. In the mutant, however, they divided less. This led to the lack of differentiation of the optic lobes and the reduced size of the thoracic neuromeres (*b*). Growth of the imaginal discs was also poor in these animals. Furthermore, the larval stages were prolonged and the animals died as larvae or poorly contracted pupae. Preliminary data (not shown) indicate that these defects may stem from a malfunctioning of the larval brain or ring gland, or from both. All of these defects were progressively improved if the embryos had been supplied with increased amounts of maternal *sgg* gene product. *c*, Schematic drawings (derived from camera-lucida drawn shapes) of third-instar larval nervous systems (lateral view) of mutant animals derived from mothers carrying 1, 2, 3 or 4 doses of *sgg*⁺, from top to bottom. The central nervous system shown at the bottom was from a wild-type sibling. Gross morphological inspection showed improved growth, notably of the optic lobes and thoracic neuromeres. We interpret



these results to mean that increasing amounts of maternally-supplied *sgg* gene product can lead to a more functional larval brain, better able to support further larval growth.

METHODS. Females were of the following genotypes: (1) *FM6/sgg*^{D127} *w* *s*; (2) *C(1)DX/Y*; (3) *C(1)DX/Y*; *dup(1;2)w*^{+70h}, *sgg*^{+/+}; (4) *C(1)DX/Y*; *dup(1;2)w*^{+70h}, *sgg*^{+/+}; *dup(1;3)w*^{+67k27}, *sgg*^{+/+}. These were crossed to *sgg*^{D127} *w* *s/w*⁺ *Y*, *sgg*⁺ males. The resulting mutant zygotes (*w*) could be recognized by their white malpighian tubules. *a*, *b*, Larval nervous systems from female and mutant male siblings from cross with females of genotype 2. Labelling of dividing cells and their progeny was as described previously²⁹. Larvae were fed for three days with BUDR, and larval central nervous systems were then dissected; detection of BUDR was by immunocytochemical staining using a peroxidase conjugate.

Staining with neuronal markers revealed neural hyperplasia (Fig. 2). This phenotype is not unlike mutants of the neurogenic class. The embryos could differentiate to give cuticle but were short. Increased amounts of zygotically derived *sgg*-gene product revealed cuticular defects characteristic of segmentation mutants, and could lead to an almost complete rescue of the pattern (Fig. 3). The maternally rescued embryos went through embryogenesis normally but died as defective larvae with abnormal brains (Fig. 4). These results indicate that maternally and zygotically derived gene-products must perform similar functions. Furthermore, there are formal similarities between the embryonic and imaginal phenotypes of mutants for *sgg*. The *sgg* mutation causes hyperplasia of the peripheral nervous system—that is, vast numbers of densely packed sensory bristles differentiate over the entire adult epidermis^{15,24}. This phenotype resembles that of gain-of-function mutants of the achaete-scute complex²⁵.

Similarities are thus seen between different phenotypes of *sgg* mutants and the phenotypes of mutants of both the pro-neural and neurogenic classes. How, therefore, can the null phenotype of *sgg* be explained in terms of the known functions of the pro-neural and neurogenic gene classes? One possibility is that the *sgg*-gene product may normally serve to repress one or more of the pro-neural genes that are required for the specification of neural cells. Such target genes could be those of the achaete-scute complex. Analysis of the phenotypes of embryos doubly mutant for *achaete-scute* and *sgg* (data not shown), and gene-dose studies in the adult¹⁵, however, demonstrate that *sgg* causes neither maternal nor zygotic derepression of genes of the achaete-scute complex. Alternatively, genes of the neurogenic class encode functions that are required for intercellular communication²⁰. The requirement of *sgg* for normal bristle-cell segregation¹⁵ and the observed cell-autonomy of the *sgg* mutant phenotype²⁴ suggests that *sgg*-mutant cells are independent of the inhibitory signal. A role of *sgg* for lateral inhibition is consistent with the neural hyperplasia exhibited by the zygotically rescued embryos. Furthermore, the requirement for *sgg* in the segmentation process could also reflect altered

cell communication.

In view of the different phenotypes observed, *sgg* probably interferes with some general cell property; in the absence of this property cells become neural by default rather than as the result of a specific neural regulatory switch. Although they are not defective in an essential component of the cell membrane, *sgg* mutant cells might be unable to establish normal cellular contacts. □

Received 1 July; accepted 25 August 1989.

1. Campos-Ortega, J. A. *Trends Neurosci.* **8**, 245–250 (1985).
2. Hartenstein, V. & Campos-Ortega, J. A. *Wilhelm Roux Arch. dev. Biol.* **193**, 308–325 (1984).
3. Ghysen, A. & Dambly-Chaudière, C. *Gif Lectures in Neurobiology* (ed. Masson) (*J. Physiol., Paris*) (in press).
4. Ghysen, A. & Dambly-Chaudière, C. *Genes Dev.* **2**, 495–501 (1988).
5. Garcia-Bellido, A. & Santamaria, P. *Genetics* **88**, 469–486 (1978).
6. Dambly-Chaudière, C. & Ghysen, A. *Genes Dev.* **1**, 297–306 (1987).
7. Jimenez, F. & Campos-Ortega, J. A. *Neurogenetics* **4**, 179–200 (1987).
8. Cabrera, C. V., Martinez-Arias, A. & Bate, M. *Cell* **50**, 425–433 (1987).
9. Romani, S., Campuzano, S. & Modolelli, J. *EMBO J.* **6**, 2085–2092 (1987).
10. Sulston, J. E. & White, J. G. *Dev. Biol.* **78**, 577–597 (1980).
11. Doe, C. Q. & Goodman, C. S. *Dev. Biol.* **111**, 206–219 (1985).
12. Taghert, P. H., Doe, C. Q. & Goodman, C. S. *Nature* **307**, 163–165 (1984).
13. Wigglesworth, V. B. *J. expl. Zool.* **17**, 180–220 (1940).
14. Richelle, J. & Ghysen, A. *Dev. Biol.* **70**, 418–437 (1979).
15. Simpson, P. & Carteret, C. *Development* **106**, 57–66 (1989).
16. Lehman, R., Dietrich, V., Jimenez, F. & Campos-Ortega, J. A. *Wilhelm Roux Arch. dev. Biol.* **190**, 226–229 (1981).
17. Lehmann, R., Jimenez, F., Dietrich, U. & Campos-Ortega, J. A. *Wilhelm Roux Arch. dev. Biol.* **192**, 62–74 (1983).
18. Dietrich, U. & Campos-Ortega, J. A. *J. Neurogenetics* **1**, 315–337 (1984).
19. Hartenstein, V. & Campos-Ortega, J. *Wilhelm Roux Arch. dev. Biol.* **195**, 210–221 (1986).
20. Campos-Ortega, J. *Trends Neurosci.* **11**, 400–405 (1988).
21. Jan, L. Y. & Jan, Y. N. *Proc. natl Acad. Sci. USA* **79**, 2700–2704 (1982).
22. Bier, E., Ackerman, L., Barbel, S., Jan, L. Y. & Jan, Y. N. *Science* **240**, 913–916 (1988).
23. Ripoll, P., El Messal, M., Laran, E. & Simpson, P. *Development* **103**, 757–767 (1988).
24. Simpson, P., El Messal, M., Moscoso del Prado, J. & Ripoll, P. *Development* **103**, 391–401 (1988).
25. Garcia-Alonso, L. & Garcia-Bellido, A. *Wilhelm Roux Arch. dev. Biol.* **195**, 259–264 (1986).
26. Bussón, D., Gans, M., Komitopoulou, K. & Masson, M. *Genetics* **105**, 309–325 (1983).
27. Lehman, R. & Nüsslein-Volhard, C. *Dev. Biol.* **119**, 402–417 (1987).
28. Jurgens, G., Wieschaus, E., Nüsslein-Volhard, C. & Kluding, H. *Wilhelm Roux Arch. dev. Biol.* **193**, 283–295 (1984).
29. Truman, J. W. & Bate, M. *Dev. Biol.* **125**, 145–157 (1988).

ACKNOWLEDGEMENTS. We thank Cathie Carteret and Claudine Ackerman for technical assistance, Yves Grau for comments on the manuscript and permission to cite unpublished results, Y. N. Yan for providing the monoclonal antibody Mab44C11 and Pedro Ripoll for discussions. This work was funded by the CNRS and INSERM.

Restoration of normal function in genetically defective myotubes by spontaneous fusion with fibroblasts

Nirupa Chaudhari*, Rona Delay† & Kurt G. Beam*

* Department of Physiology and † Department of Anatomy and Neurobiology, Colorado State University, Fort Collins, Colorado 80523, USA

MUSCULAR dysgenesis in mice is a genetic disease of skeletal muscle caused by the recessive mutation *mdg* (ref. 1). Muscle fibres in affected mice are paralysed because of the failure of excitation-contraction coupling^{2,3}. Unlike normal myotubes in primary culture, dysgenic myotubes do not contract, either spontaneously or in response to electrical stimulation. The deficiency results from mutation of the gene for the skeletal muscle dihydropyridine receptor⁴, an essential sarcolemmal component both of excitation-contraction coupling and of the slow calcium-ion channel^{4,5}. It has recently been shown⁶ that the addition of fibroblasts from normal (but not dysgenic) mice to cultures of dysgenic myotubes can restore spontaneous contractions in a small fraction of these myotubes, but the mechanism for this 'rescue' was not determined. In principle⁷, if fibroblast nuclei were able to incorporate into myotubes, such nuclei could then supply the missing muscle-specific gene product. We have now investigated this possibility using nuclear, cytoplasmic and plasmalemmal markers. We report that the rescue to contractile ability in genetically paralysed dysgenic muscle is mediated by the previously unrecognized ability of fibroblasts to fuse spontaneously with developing myotubes.

We observed the functional rescue of dysgenic myotubes by their co-culture with cloned dermal fibroblasts from normal mice and rats, and with several fibroblast lines including 3T3, Rat2 and NRK-52E (derived from mouse embryos, rat embryos and rat kidney, respectively). The incidence of rescue varied from 2 to 28% of all myotubes, depending on the assay (spontaneous versus electrically evoked contractions), the identity of the added fibroblasts, and the time of addition of fibroblasts relative to myotube development. The addition of normal fibroblasts during myoblast proliferation was most effective; very little rescue occurred if fibroblasts were added after myotube development. Living fibroblasts seemed to be essential, because neither a medium that was conditioned by normal fibroblasts⁶, nor a substrate of killed (water-lysed or ultraviolet-irradiated) normal fibroblasts, produced rescue.

To test our hypothesis that rescue reflected the fusion of fibroblasts with forming myotubes, we initially used a clonal line of 3T3 fibroblasts (CIW), which contained a stably integrated copy of the *Escherichia coli* gene for β -galactosidase⁸. The cytoplasm of CIW cells reacted positively when stained for β -galactosidase (Fig. 1a). When CIW cells were added to cul-

tures of rat myoblasts or phenotypically normal (+/*mdg*?) mouse myoblasts, a small fraction of the myotubes that developed clearly contained β -galactosidase in their cytoplasm (Fig. 1b). In two such experiments, 1.5% of 1,600 rat myotubes and 6.0% of 650 normal (+/*mdg*?) mouse myotubes stained bright blue. Neither medium conditioned by CIW cells, nor homogenates of CIW cells, induced β -galactosidase activity in exposed myotubes. Because β -galactosidase is a protein of relative molecular mass 116,000 (M_r , 116K), it is not expected to pass through gap junctions, which have been reported to form between myotubes and non-muscle cells in culture^{2,9}. The presence of β -galactosidase within myotubes, therefore, indicated that fibroblasts are able to fuse with myotubes developing in culture.

To test whether fusion of fibroblasts is necessary for the rescue of dysgenic myotubes, we added CIW cells to dysgenic myoblast cultures. Rescued dysgenic myotubes were identified by their ability to twitch in response to electrical stimulation (Fig. 1c). In two separate experiments, most of the rescued myotubes were stained pale to bright blue (Fig. 1d and Table 1). The absence of staining in 12–15% of rescued myotubes (see Table 1) implies that either β -galactosidase activity was below the level of detection in these rescued myotubes or that fusion was not required for their rescue. Thus, we established an independent assay for fusion: after staining with Hoechst 33342, mouse nuclei displayed a punctate fluorescence (Fig. 2b), whereas rat nuclei displayed a more homogeneous fluorescence (Fig. 2b, indicated by the arrow). After addition of normal rat fibroblasts to dysgenic myoblast cultures, rescued myotubes were identified by the presence of spontaneous or electrically evoked contractions. Every rescued myotube contained at least one normal rat nucleus (Fig. 2b, d and Table 1). To confirm that the rat nucleus was located inside the dysgenic myotube (rather than in a rat fibroblast adhering to the myotube), the cultures were stained with rat-specific monoclonal antibody against Thy-1.1, an antigen expressed on the surface of fibroblasts¹⁰. Immunostaining with this antibody confirmed that rat nuclei visually identified as being incorporated into rescued myotubes were always located within the myotubes, rather than in overlying fibroblasts (Fig. 2e).

Electron microscope examination of dysgenic myotubes rescued by rat fibroblasts further showed that rat nuclei were located within the sarcolemma. Such nuclei possessed an intact nuclear envelope with no evidence of fragments or invaginations of the sarcolemma in their vicinity. The electron micrographs also showed that portions of the rescued myotubes that were distant from the incorporated rat nucleus had the ultrastructural appearance of typical dysgenic myotubes¹¹. For instance, myofilaments were not ordered into sarcomeres but were clumped throughout the myoplasm (Fig. 3c). In rescued myotubes, the organization of myofilaments into a normal sarcomeric pattern was observed, but only in the neighbourhood of the rat nucleus (Fig. 3d). Similar results were observed for two rescued myotubes additionally examined.

TABLE 1 Rescue of dysgenic myotubes by fusion with fibroblasts

Experiment	Type of fibroblast added	Method of assay for fusion	Frequency of marker indicating fusion (%)	
			Rescued myotubes	Non-rescued myotubes
A	CIW	β -galactosidase	85 (n=13)	20 (n=300)
B	CIW	β -galactosidase	88 (n=16)	29 (n=250)
C	Rat dermal	Nuclear morphology	100 (n=17)	31 (n=74)
D	NRK-52E	Nuclear morphology	100 (n=10)	30 (n=56)

Fusion of dysgenic myotubes with fibroblasts is required for rescue. Primary cultures of dysgenic myoblasts were co-cultured with fibroblasts as described for CIW cells in Fig. 1 legend. Primary dermal fibroblasts were cultured and cloned in our laboratory; NRK-52E fibroblasts were obtained from American Type Culture Collection, Rockville. Myotubes were tested for rescue by electrical stimulation (experiments A, B) or by observing spontaneous contractions (experiments C, D). In experiment A, myotubes were stimulated early (day 8), before any spontaneous contractions had begun in the culture. In experiment B, myotubes were grown in the presence of 10 μ M tetrodotoxin to inhibit spontaneous contractions. Cultures were stained for β -galactosidase (experiments A, B) or with Hoechst 33342 (experiments C, D), as described in the legends to Figs 1 and 2, respectively.

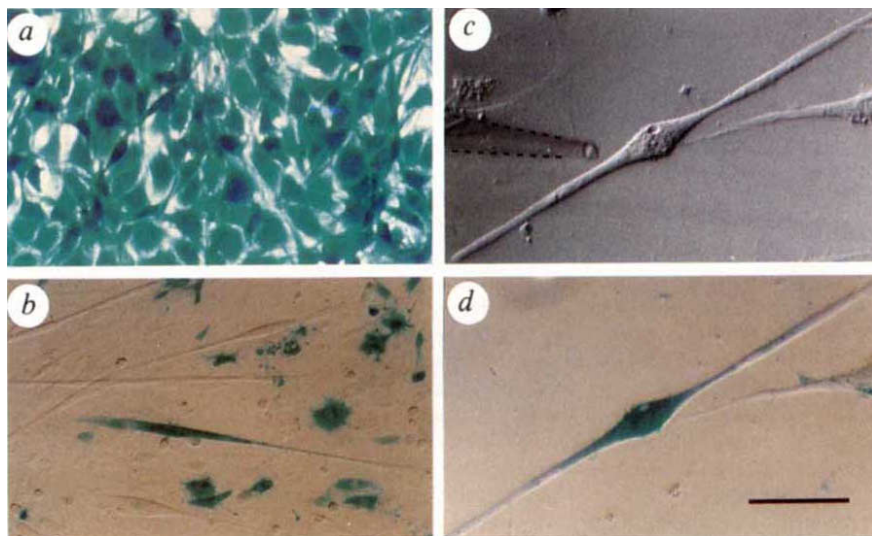
Because co-culture with normal neurons (that is, dissociated spinal-cord cells) can restore contractile ability to dysgenic myotubes¹², it has been suggested that the primary alteration in muscular dysgenesis is extra-muscular (for example, neuronal)¹³. But the rescue obtained with neuronal co-culture could be explained by the fusion of non-muscle cells with

myotubes, especially as fibroblasts and other cell types contaminate preparations of dissociated spinal neurons.

As in normal muscle, the rescued dysgenic myotubes contracted in response to electrical stimulation even under conditions that prevented the entry of calcium through voltage-gated Ca^{2+} channels (Fig. 4). Together with the observation of normal

FIG. 1 Naturally occurring fusion of fibroblasts with myotubes as indicated by β -galactosidase staining. *a*, CIW is a fibroblast line of cells which contain a stably integrated bacterial gene for β -galactosidase. After fixation and histochemical staining for this enzyme, all CIW cells display a blue reaction product. *b*, CIW cells spontaneously fuse with normal myotubes. CIW fibroblasts were added to a normal rat myoblast culture one day after initial plating. Four days later, the culture was fixed and stained for β -galactosidase activity. In this experiment, ~1.5% of all myotubes had brightly stained cytoplasm, as shown. *c*, Addition of CIW fibroblasts leads to recovery of contractile ability in some dysgenic myotubes. CIW cells were added to a one-day-old dysgenic myoblast culture. On day 8 of culture, individual myotubes were tested for their ability to contract in response to electrical stimulation using an agar-plugged extracellular pipette⁴. Myotubes that responded with a normal contraction were photographed for later identification. *d*, Most rescued dysgenic myotubes clearly displayed the blue staining, which indicates that a CIW cell had fused with the myotube. After stimulation and photography for identification of rescued myotubes, cultures were fixed and stained for β -galactosidase. The myotube shown is the same as that in *c*. Scale bar: 100 μm for *a*, *c* and *d*; 200 μm for *b*.

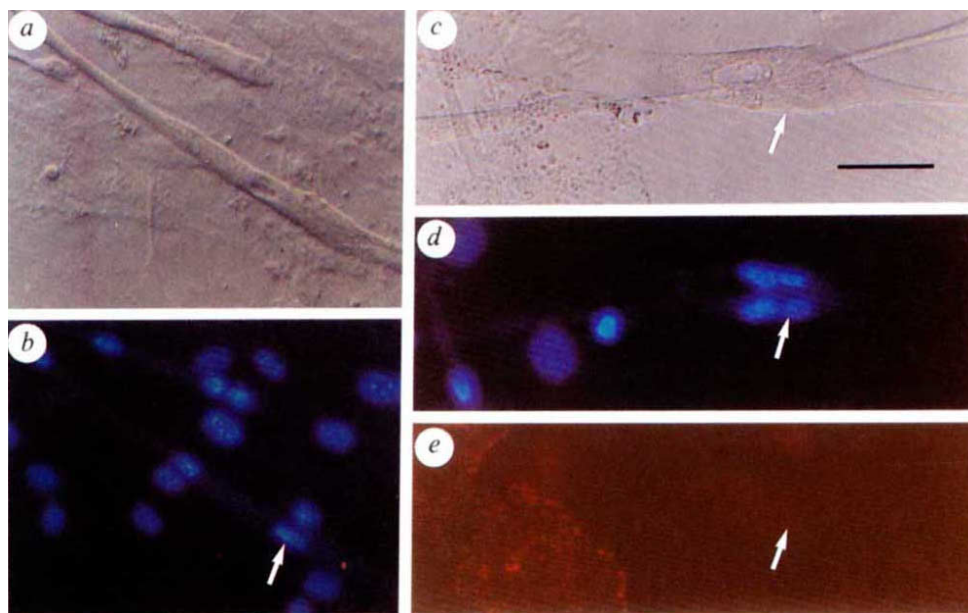
METHODS. Myoblasts, prepared from neonatal dysgenic mice, normal mice or normal rats, were plated (at a density of $\sim 10^5$ per 35 mm dish) and cultured as described previously¹⁷. One day after initial plating of the myoblasts, CIW fibroblasts were added (10^4 per dish). After 2–3 days of co-culture, the serum content of the medium was changed (from 20% fetal calf to 10% horse), and excess fibroblasts were removed by treatment for 1–2 days with 10 μM cytosine arabinoside. Four to ten days after the initial



plating, normal cultures were fixed and stained for β -galactosidase as described⁸. Some dysgenic cultures were maintained in the continual presence of 10 μM tetrodotoxin to prevent spontaneous contraction in the rescued myotubes, because we observed that extensive contractile activity in normal or rescued dysgenic myotubes leads to a progressive loss of β -galactosidase activity. The tetrodotoxin-treated dysgenic cultures were tested for contractile ability 9 days after initial plating; the untreated dysgenic cultures were tested 6–8 days after plating. Cultures were tested for the presence of rescued myotubes by focal electrical stimulation as described⁴ and then were fixed and stained for β -galactosidase.

FIG. 2 Dysgenic myotubes rescued by co-culture with fibroblasts contain incorporated fibroblast nuclei. Rat2 cells were added to a dysgenic myoblast culture, and myotubes that exhibited spontaneous contractions were identified 8 days later and photographed using Hoffman differential contrast optics; two such rescued myotubes are shown (*a*, *c*). The same myotubes, viewed under ultraviolet fluorescence after staining with Hoechst 33342, are illustrated in *b* and *d*, respectively. Each contained an incorporated rat nucleus (indicated by an arrow), which is distinguishable by its homogeneously stained chromatin. Mouse nuclei display a punctate chromatin staining. In *e*, the same myotube as in *c* and *d* is shown after immunostaining for rat Thy-1.1 antigen. Note that the two rat fibroblasts (that align with the two large homogeneously stained nuclei in the left-hand portion of *d*), display red labelling for the presence of Thy-1.1 in their plasma membranes. By contrast, the rat nucleus indicated by the arrow is not surrounded by red staining, indicating that this nucleus is incorporated into the myotube and is no longer part of a fibroblast. Scale bar, 50 μm for all parts of figure.

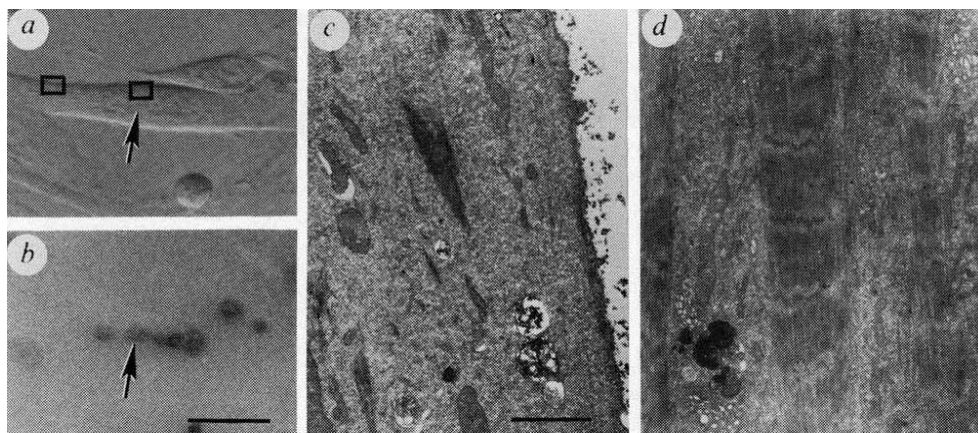
METHODS. Rat2 fibroblasts (obtained from American Type Culture Collection) were added to cultures of dysgenic myoblasts as described in Fig. 1 legend. Myotubes that twitched spontaneously were identified by marks on the lower surface of the culture dish. Living cultures were stained with Hoechst 33342 at 3 $\mu\text{g ml}^{-1}$ for 10 min at 37 $^{\circ}\text{C}$ in culture medium. For immunocytochemical



staining, cultures were incubated at 37 $^{\circ}\text{C}$ for 10 min with a mouse monoclonal antibody against rat Thy-1.1 antigen (Chemicon, diluted 100-fold in culture medium). After washing, a tetramethylrhodamine isothiocyanate-conjugated goat anti-mouse IgG (Sigma, diluted 50-fold in culture medium) was applied for 10 min at 37 $^{\circ}\text{C}$. After several washes to remove unbound antibody, cultures were fixed in 4% paraformaldehyde in phosphate-buffered saline.

FIG. 3 Ultrastructural restoration only occurs near the normal nucleus within a rescued dysgenic myotube. A dysgenic myotube, rescued by the addition of rat NRK-52E fibroblasts, was viewed by Hoffman differential contrast optics (a) and by ultraviolet fluorescence after staining with Hoechst 33342 (b). The rescued myotube contains a single rat nucleus (indicated by the arrow). Electron micrographs of the regions distant from the rat nucleus show disorganized myofilaments (c) typical of non-rescued dysgenic myotubes, whereas those of the region near the rescuing rat nucleus show sarcomeric organization of the contractile proteins (d). The regions illustrated in c and d correspond to those in the left and right boxes in a, respectively. Scale bars, either 50 μm (a, b) or 2 μm (c, d).

METHODS. NRK-52E fibroblasts were added to a dysgenic myoblast culture. After 9 days, a myotube that twitched spontaneously was identified, stained with Hoechst 33342 and photographed. The culture was then fixed in 2.5%



glutaraldehyde, 0.5% tannic acid in 0.1 M phosphate buffer, pH 7.4. The culture was post-fixed in 2% osmium, stained with uranyl acetate and embedded in plastic. Thin sections were post-stained with uranyl acetate and Sato's lead stain¹⁸, and were viewed with an electron microscope (JEOL 1200EX).

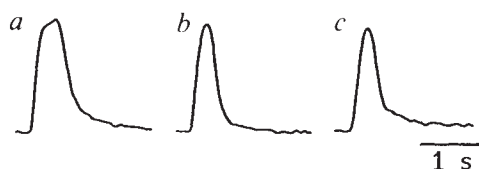


FIG. 4 Calcium entry is not required for electrically evoked contractions of dysgenic myotubes rescued by fibroblasts. Contractions were elicited with the rescued myotube bathed in normal physiological saline (a), in Ca^{2+} -free saline (b) and after return to normal saline (c). Tetrodotoxin (3 μM) was present to block spontaneous contractions. The normal saline¹⁷ contained 2 mM Ca^{2+} and 1 mM Mg; the Ca^{2+} -free saline contained 3 mM Mg^{2+} . Electrical stimulation and optical recording of contractions were as described previously⁴. Rescue of this dysgenic myotube was effected by addition of Rat2 fibroblasts.

dihydropyridine-sensitive Ca^{2+} currents in dysgenic myotubes that have been rescued by fibroblasts⁶, this indicates that a normal gene for the skeletal muscle dihydropyridine receptor is expressed in these myotubes. Because the genome of dysgenic mice lacks a functional copy of the normal gene for the dihydropyridine receptor⁴, we presume that muscle cytoplasm induces the fibroblast nucleus to express muscle-specific genes, as has been demonstrated through artificially induced fusion⁷.

We have shown here that in tissue culture, fibroblasts are able to fuse with developing myotubes. Fibroblasts seem to fuse very rarely or not at all if they are added after myotubes have fully formed. We do not yet know whether a similar fusion process also can occur *in vivo* or, conversely, whether mechanisms exist to prevent such fusions. If fibroblasts can fuse with developing muscle *in vivo*, they might serve as a useful donor cell-type for the treatment of genetic diseases of muscle in humans¹⁴⁻¹⁶. □

14. Law, P. K., Goodwin, T. G. & Li, H.-J. *Transplantation Proc.* **20**, 1114-1119 (1988).
15. Law, P. K., Goodwin, T. G. & Wang, M. G. *Muscle & Nerve* **11**, 525-533 (1988).
16. Partridge, T. A., Morgan, J. E., Coulton, G. R., Hoffman, E. P. & Kunkel, L. M. *Nature* **337**, 176-177 (1989).
17. Beam, K. G., Knudson, C. M. & Powell, J. A. *Nature* **320**, 168-170 (1986).
18. Sato, T. *J. Electron. Microsc., Tokyo* **17**, 158-159 (1968).

ACKNOWLEDGEMENTS. We thank Dr J. Sanes for suggesting *lacZ* as a marker and providing CIW cells, Dr J. R. Fallon for suggesting Thy-1.1 as a marker, and many colleagues for comments on the manuscripts. This research was supported by the NIH and the Muscular Dystrophy Association (K.G.B.).

Recognition of cluster of differentiation 1 antigens by human $\text{CD4}^- \text{CD8}^-$ cytolytic T lymphocytes

Steven Porcelli*, Michael B. Brenner*, Julia L. Greenstein†, Steven P. Balk‡§, Cox Terhorst‡|| & Paul A. Bleicher‡¶

*Laboratory of Immunochimistry, Dana-Farber Cancer Institute, and Department of Rheumatology and Immunology, Brigham and Women's Hospital; †Laboratory of Molecular Immunology, Dana-Farber Cancer Institute; Departments of ‡Dermatology and ||Pathology, Harvard Medical School; §Department of Hematology and Oncology, Beth Israel Hospital; Boston, Massachusetts 02115, USA
† ImmuLogic Pharmaceutical Corp, Cambridge, Massachusetts 02139, USA

HUMAN cluster-of-differentiation 1 (CD1) is a family of cell surface glycoproteins of unknown function expressed on immature thymocytes, epidermal Langerhans cells and a subset of B lymphocytes¹⁻³. Three homologous proteins, CD1a, b and c, have been defined serologically⁴, and the CD1 gene locus on human chromosome 1 contains five potential CD1 genes⁵. Analysis of the predicted amino-acid sequences of CD1 molecules reveals a low but significant level of homology to major histocompatibility complex (MHC) class I and class II molecules^{5,6}, and, like MHC class I molecules, CD1 molecules are associated non-covalently with β_2 -microglobulin⁷. These structural similarities to known antigen-presenting molecules, together with the expression of CD1 on cells capable of antigen presentation, suggest a role for CD1 molecules in antigen recognition by T cells. Here we demonstrate the specific recognition of CD1a by a $\text{CD4}^- \text{CD8}^- \alpha\beta$ T-cell receptor (TCR) expressing cytolytic T lymphocyte (CTL) line and the specific

Received 10 April; accepted 29 August 1989.

1. Pai, A. C. *Dev. Biol.* **11**, 82-92 (1965).
2. Powell, J. A. & Fambrough, D. M. *J. cell Physiol.* **82**, 21-38 (1973).
3. Klaus, M. M., Scordilis, S. P., Rapalus, J. M., Briggs, R. T. & Powell, J. A. *Dev. Biol.* **99**, 152-165 (1983).
4. Tanabe, T., Beam, K. G., Powell, J. A. & Numa, S. *Nature* **336**, 134-139 (1988).
5. Rios, E. & Brum, G. *Nature* **325**, 717-720 (1987).
6. Courbin, P., Koenig, J., Ressouches, A., Beam, K. G. & Powell, J. A. *Neuron* **2**, 1341-1350 (1989).
7. Blau, H. M. *et al. Science* **230**, 758-766 (1985).
8. Sanes, J. R., Rubenstein, J. L. R. & Nicolas, J.-F. *EMBO J.* **5**, 3133-3142 (1985).
9. Bonner, P. H. *Dev. Brain Res.* **38**, 233-244 (1988).
10. Stern, P. L. *Nature new Biol.* **246**, 76-78 (1973).
11. Barker, B. Q. *J. Neuropathol. exp. Neurol.* **36**, 100-127 (1977).
12. Koenig, J., Bournaud, R., Powell, J. A. & Rieger, F. *Dev. Biol.* **92**, 188-196 (1982).
13. Rieger, F. *et al. Nature* **330**, 563-566 (1987).

recognition of CD1c by a CD4⁺CD8⁺ $\gamma\delta$ TCR CTL line. The interaction of CD1-specific CTLs with CD1⁺ target cells appeared to involve the CD3-TCR complex, and did not show evidence of MHC restriction. These results suggest that for a subset of T cells, CD1 molecules serve a function analogous to that of MHC class I and II molecules.

We have previously described a CD4⁺CD8⁺ $\gamma\delta$ TCR CTL line, IDP2, which showed specific MHC-independent lysis of the MOLT-4 thymic leukaemia-cell line⁸. By screening another 21 randomly derived CD4⁺CD8⁺ peripheral-blood T-cell lines (13 expressing $\gamma\delta$ TCR, 8 expressing $\alpha\beta$ TCR), we found one other CTL, BK6 ($\alpha\beta$ TCR), with a similar pattern of target recognition. As MOLT-4 expresses CD1a, b and c, we considered that one of these could be the specific target structure recognized by these CTL and examined this by testing the ability of anti-CD1a, b or c monoclonal antibodies to block lysis. As seen in Fig. 1, lysis of MOLT-4 by BK6 was inhibited by the three anti-CD1a monoclonal antibodies tested, but not by those specific for CD1b or CD1c. Conversely, lysis by IDP2 was inhibited by the four anti-CD1c monoclonal antibodies tested, but not by those specific for CD1a or CD1b (Fig. 1). Anti-MHC antibodies specific for either monomorphic determinants of class I (W6/32 and BB7.7, Fig. 1) or class II (LB3.1, data not shown) had no effect on the lysis of MOLT-4 cells by either cell line. Anti-CD8 monoclonal antibody OKT8, which does not bind detectably to either CTL, also had no effect (Fig. 1).

Similar blocking experiments were carried out with monoclonal antibodies directed against components of the CD3-associated TCR complex (Fig. 2). Anti-CD3 antibodies (SPV-T3b) inhibited target-cell lysis by both T-cell clones, supporting a role for the CD3-TCR complex in CD1 recognition. The effects of antibodies against other parts of the CD3-TCR complex supported this idea, as the lysis of MOLT-4 by BK6 was inhibited by an antibody (BMA031) that binds specifically to the $\alpha\beta$ TCR complex, whereas lysis by IDP2 was inhibited by antibodies that bind to the δ TCR (anti-TCR δ 1) or γ TCR (anti-Ti γ A; ref. 9) chains.

To demonstrate directly that CTLs BK6 and IDP2 recognized CD1a and CD1c respectively, we generated a series of CD1 transfectant cell lines. Transfectants of murine T-T hybridoma BY155.16 (ref. 10) and murine L-cell fibroblasts expressing CD1a or CD1c were used as targets. As predicted from the monoclonal antibody blocking data, a CD1a transfectant of BY155.16 was lysed by BK6, whereas a CD1c transfectant was not (Fig. 3a, left). Conversely, IDP2 lysed the CD1c transfectant but not the CD1a transfectant (Fig. 3a, right). Similar results were obtained using L cells transfected with CD1a and CD1c as targets (Fig. 3b). In all cases, monoclonal antibodies specific for the appropriate form of CD1 blocked specific lysis of the CD1 transfectants (Fig. 3a-c). Recognition of the appropriate form of CD1, therefore, could be demonstrated for each clone, regardless of the cell type into which CD1 was introduced.

Lysis of murine transfectants was relatively inefficient, requiring high effector-to-target-cell ratios to achieve significant levels (Fig. 3a, b). Such inefficiency may have been due to the absence from murine cells of ligands (for example, ICAM-1 (intracellular adhesion molecule-1) and LFA-3 (lymphocyte-function associated antigen-3) for human T-cell accessory molecules important in CTL-target interactions¹¹. This possibility was supported by results obtained with CD1 transfectants of the human rhabdomyosarcoma-cell line RD (ref. 12). BK6 lysed a CD1a⁺ transfectant of the human cell line much more efficiently than the corresponding CD1a⁺ murine transfectants (Fig. 3c), despite a lower level of CD1a expression on the human cells (data not shown). Lysis of the CD1a⁺ RD transfectant by BK6 was specific for CD1a, as it was blocked by anti-CD1a monoclonal antibody and no lysis of a CD1b⁺ RD transfectant was observed (Fig. 3c).

Lysis of CD1⁺ targets by CTLs BK6 and IDP2 implied that the CD1-dependent interaction between effector and target cells resulted in activation of the CTLs. To confirm this by an indepen-

dent approach, we examined the production of interleukin-2 (IL-2) by BK6 after incubation with CD1a⁺ target cells. Culturing BK6 in the presence of the CD1a⁺ RD transfectant cells caused measurable IL-2 secretion, which was not seen if BK6 was cultured under the same conditions with the CD1b⁺ transfectant (Fig. 4). The IL-2 secretion stimulated by CD1a recognition was completely inhibited by a monoclonal antibody specific for CD1a, whereas IL-2 secretion stimulated by phytohaemagglutinin was not affected by the addition of this anti-CD1a antibody (Fig. 4).

Our results show that some peripheral T cells can recognize CD1 molecules, and that this recognition triggers cellular events

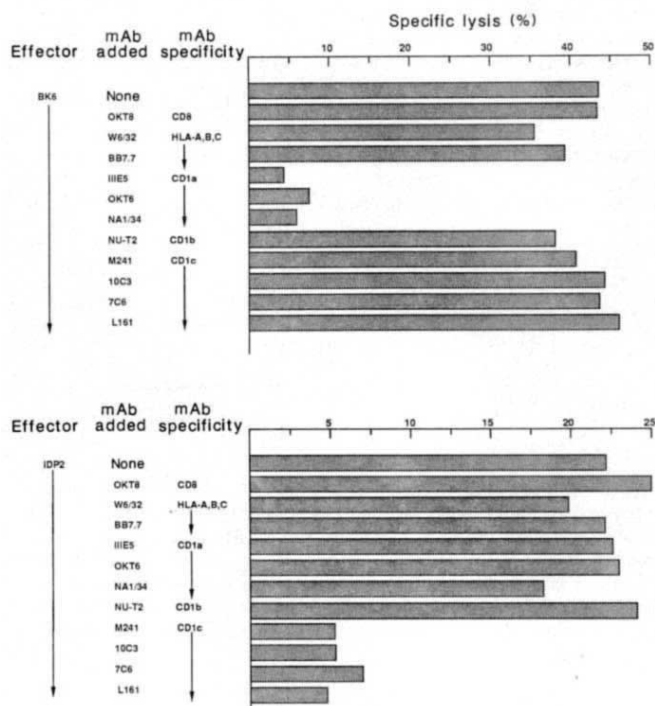


FIG. 1 Anti-CD1 monoclonal antibodies (mAb) block lysis of MOLT-4 cells by CTLs BK6 and IDP2. MOLT-4 T-leukaemia cell line is CD1abc⁺, MHC class I⁺/II⁺, CD3⁺CD4⁺CD8⁺. Lysis by $\alpha\beta$ TCR CTL BK6 (effector:target [E:T] ratio 20:1, top panel) was inhibited only by anti-CD1a antibodies (III E5 (ref. 20), OKT6 (ref. 21) and NA1/34 (ref. 1)). Lysis by $\gamma\delta$ TCR CTL IDP2 (E:T ratio 10:1, bottom panel) was inhibited only by anti-CD1c antibodies (M241 (ref. 22), 10C3 (ref. 23), 7C6 (ref. 23) and L161 (ref. 4)). Anti-MHC class I antibodies W6/32 (ref. 24) and BB7.7 (ref. 24), anti-CD1b antibody NU-T2 (ref. 4) and anti-CD8 antibody OKT8 (ref. 21) did not inhibit lysis. Results are means of triplicate values with standard errors of less than 8%, and are representative of three separate experiments.

METHODS. Assay of chromium release and calculation of per cent specific lysis have been described⁸. ⁵¹Cr-labelled MOLT-4 target cells were incubated with the indicated monoclonal antibody (10⁻³ dilution of ascites fluid) for 30 min at room temperature before addition of CTLs. All antibodies bound to saturation under these conditions. The derivation and characterization of the IDP2 line have been described⁸. CTL line BK6 was derived from the peripheral blood of a patient with systemic lupus erythematosus. Briefly, mononuclear cells recovered after treatment with anti-CD4 and -CD8 monoclonal antibodies and rabbit complement (Cedar Lane) were expanded for two weeks in IL-2-containing medium⁸ with anti-CD3 antibody SPV-T3b (ref. 25) ascites at a 10⁻⁴ dilution as a polyclonal T-cell mitogen. Cells were then stained with a mixture of monoclonal antibodies OKT4 (ref. 21) (anti-CD4), OKT8 and anti-TCR δ 1 (ref. 26), and unstained cells isolated by sorting with an Epics IV FACS (Coulter Electronics). BK6 was obtained by seeding these cells at limiting dilution (2.5 cells per well, plating efficiency 5%). This cell line has been in continuous culture for one year, and maintains the stable phenotype $\alpha\beta$ TCR CD4⁺CD8⁺ based on staining with monoclonal antibodies BMA031 (ref. 27) (anti- $\alpha\beta$ TCR), OKT4 and OKT8. Both IDP2 and BK6 were negative for cell-surface expression of CD1a, b and c as judged by staining with specific antibodies.

similar to those seen when T cells interact with antigens presented by self-MHC molecules. The two CTLs studied were each specific for a single form of CD1, either CD1a or CD1c. This specificity, together with the blocking by monoclonal antibodies against components of the CD3-TCR complex, indicates that the recognition is mediated by the specific antigen receptors of these CTLs. It is unlikely that this recognition of CD1 involves classical MHC molecules, as it occurred when the appropriate form of CD1 was expressed on a variety of MHC-unrelated target cells of human or murine origin, and could not be blocked by antibodies specific for monomorphic determinants of MHC molecules. Nevertheless, we cannot exclude the possibility of co-recognition of a non-polymorphic MHC determinant not susceptible to blocking by the monoclonal antibodies used. Evidence to date indicates that CD1 molecules isolated from different individuals do not show significant polymorphism^{13,14}, so the recognition we have observed is not likely to be analogous to the frequent alloreactivity shown by T cells towards polymorphic MHC determinants. The alternative possibilities that these CD1-specific CTLs are autoreactive to unmodified self-CD1 molecules, or are recognizing an antigen produced by the target cells and presented by CD1, cannot yet be resolved.

The frequency with which peripheral T cells show CD1 reactivity has not yet been studied in detail. The finding of two examples, however, in our panel of 22 CD4⁺CD8⁺ T-cell lines derived without intentional selection for CD1 recognition, indicates that this may be a relatively common reactivity, at least within this phenotypic subset. In addition, we have recently isolated and partially characterized CD4⁺CD8⁺ T-cell lines derived from synovial tissue of rheumatoid arthritis patients. These T-cell lines are 80–90% $\gamma\delta$ TCR positive and show specific cytotoxicity of CD1c⁺-transfected target cells (S.P. and M.B.B., unpublished data). The isolation of such cells from inflamed synovium and of the CD1a-reactive CTL BK6 from the blood of a systemic lupus erythematosus patient raises the possibility of a role for CD1-reactive CD4⁺CD8⁺ T cells in autoimmune diseases. In this regard, it is of interest that CD4⁺CD8⁺ $\alpha\beta$ TCR-expressing T cells have recently been shown to be expanded in the blood of humans with systemic lupus erythematosus (ref. 15;

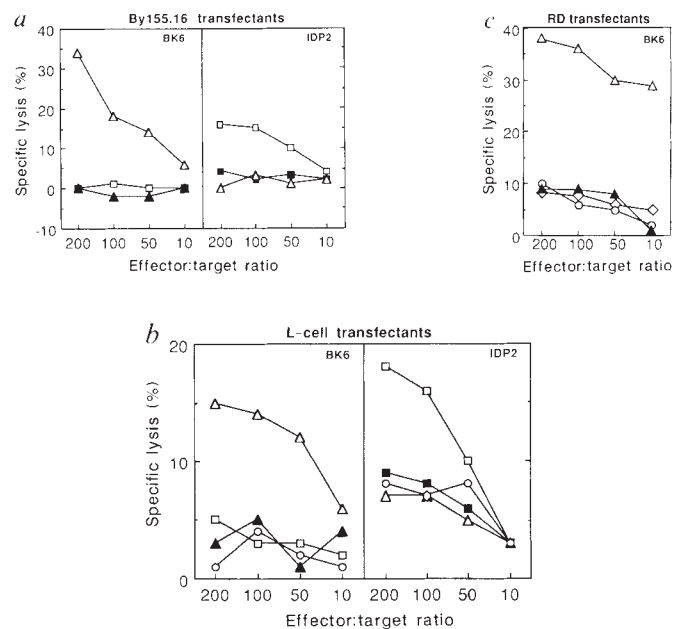


FIG. 3 CD1-specific CTLs lyse target cells transfected with the appropriate form of CD1. *a*, Lysis of BY155.16 cells transfected with CD1a or CD1c; *b*, Lysis of untransfected L cells or of L cells transfected with CD1a or CD1c; *c*, Lysis of untransfected RD cells or of RD cells transfected with CD1a or CD1c. Target cells indicated by symbols as follows: CD1a⁺ transfectants (Δ), CD1c⁺ transfectants ($\square$), CD1a⁺ transfectants preincubated with IIIE5 (anti-CD1a) antibody ($\blacktriangle$), CD1c⁺ transfectants preincubated with M241 (anti-CD1c) antibody ($\blacksquare$), untransfected target cells ($\circ$, *b* and *c* only), CD1b⁺ transfectant ($\diamond$, *c* only).

METHODS. Full-length cDNA clones for CD1a, b and c²⁸ were directionally cloned into the *XhoI*-*Bam*HI sites of the expression vector pSR α -NEO (ref. 29) using standard techniques³⁰. CD1-pSR α -NEO constructs were introduced into BY155.16 cells by electroporation³¹, and into RD and L cells by calcium phosphate mediated transfer³². Cells resistant to G418 (Gibco; 2 mg ml⁻¹ for BY155.16, 1 mg ml⁻¹ for L cells, and 0.5 mg ml⁻¹ for RD cells) were subcloned at limiting dilution. Subclones expressing CD1a, b or c were identified by surface staining with appropriate monoclonal antibodies. ⁵¹Cr-release assay and monoclonal antibody blocking were carried out as described in legend to Fig. 1.

S.P. and M.B.B., unpublished observation), as they are in the peripheral lymphoid organs of mice with hereditary autoimmune diseases that have features similar to human systemic lupus erythematosus (ref. 16). Any conclusions concerning the pathogenetic significance of CD1-reactive T cells, however, must be considered tentative, as other studies indicate that reactivity against CD1a⁺- and CD1c⁺-transfected target cells is also detectable in the CD4⁺CD8⁺ fraction of normal human peripheral blood (P.A.B. and C.T., unpublished results).

Previously we¹⁷, and others^{18,19}, have proposed that relatively non-polymorphic MHC class I-like molecules may serve as ligands for the CD3-associated TCRs of $\gamma\delta$ TCR-positive cells. Such a role has already been suggested for antigens encoded in the TI region of the MHC, which appear to be recognized by a limited number of murine $\gamma\delta$ TCR T-cell lines¹⁸. Our results indicate that CD1 molecules may also serve as specific ligands not only for some human $\gamma\delta$ TCR positive cells, which are a majority of the circulating CD4⁺CD8⁺ T cells, but also for some CD4⁺CD8⁺ $\alpha\beta$ TCR cells, which have been reported recently to be a distinct subset of circulating T cells in normal human peripheral blood¹⁵. Although we have not demonstrated directly the presentation of a foreign antigen by CD1, our results, together with the known structure and tissue distribution of

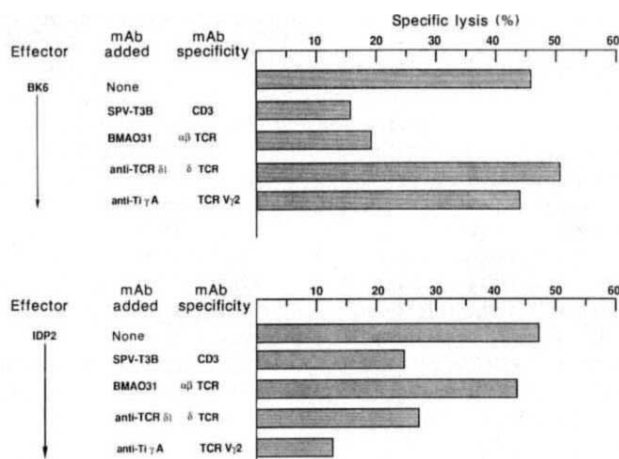


FIG. 2 Anti-TCR monoclonal antibodies block CD1-specific lysis. Lysis of MOLT-4 by CTL lines BK6 (E:T ratio 25:1) and IDP2 (E:T ratio 50:1) in the presence or absence of the indicated antibodies (mAb) was determined by ⁵¹Cr release. Results shown are representative of three separate experiments, and are means of triplicate values with standard deviations less than 10%.

METHODS. CTLs and target cells were incubated together with antibodies for 30 min at room temperature before carrying out assay of ⁵¹Cr release. BMA031 was added as purified antibody at a final concentration of 10 μ g ml⁻¹. Other antibodies were used as ascites fluid at a final dilution of 10⁻³.

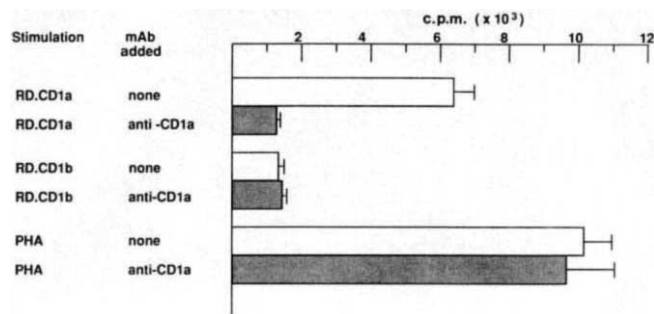


FIG. 4 Recognition of CD1a stimulates IL-2 production by CTL BK6. Culture supernatants of BK6 stimulated with CD1 transfectants of RD cells or with phytohaemagglutinin (PHA) in the presence or absence of anti-CD1a monoclonal antibody (mAb) were tested for IL-2 activity using IL-2-dependent cell line CTLL20. Results expressed as c.p.m. of [³H]thymidine incorporation by CTLL20 cells.

METHODS. BK6 cells (2.5×10^4) were cultured for 24 h in 96-well flat bottom microtitre plates (Flow Laboratories) with 10^4 RD.CD1a (CD1a⁺) or RD.CD1b (CD1b⁺) or with PHA (Gibco, 1:1,000 dilution) in 0.2 ml culture medium (RPMI-1640, 10% human serum, 2 mM L-glutamine, 1 mM HEPES, 1% penicillin-streptomycin) containing 1 ng ml^{-1} phorbol myristate acetate (PMA) (Sigma). IL-2 was assayed by a modification of a published method³³, using 4,000 CTLL20 cells and a 1:2 dilution of supernatant. Results are means ± 1 s.e.m. of triplicate cultures, and are representative of three separate experiments. [³H]thymidine incorporation by CTLL20 in the absence of IL-2 for the experiment shown was $1,158 \pm 115$ c.p.m. Similar results were obtained in the absence of PMA, although the magnitude of response was lower (data not shown).

these molecules, give rise to the hypothesis that some T cells recognize antigens presented by CD1 molecules. □

Voltage-dependent anion channels in the plasma membrane of guard cells

Bernhard U. Keller^{*†}, Rainer Hedrich^{*} & Klaus Raschke^{*}

^{*}Pflanzenphysiologisches Institut der Universität Göttingen, 34 Göttingen, FRG

[†]Max-Planck-Institut für biophysikalische Chemie, Am Fassberg 34 Göttingen, FRG.

WHEN plants assimilate CO₂ from the atmosphere, water vapour escapes. This exchange of gases can be moderated by the plant through volume changes of guard cells, which in pairs surround each stomatal pore in the epidermis and act as turgor-operated valves. When light intensity diminishes or water stress develops, guard cells shrink and stomata close. This process requires the release of ions from the guard cells. Voltage-dependent K⁺ channels^{1,2} provide a path for cations. The required corresponding passage for anions, particularly chloride and malate³, is not known at the molecular level. We have applied the patch-clamp technique⁴ to guard-cell protoplasts of *Vicia faba* and found strongly voltage-dependent activities of single anion channels in the plasma membrane (plasmalemma). These channels have a conductance of 39 pS in 100 mM KCl and are permeable for chloride and malate. They become active at membrane potentials positive to -80 mV. Their maximum activity occurs at -40 mV. The resulting outward Cl⁻ currents of individual guard cells are at least 25 pA. These voltage-dependent anion channels provide a mechanism for the control of salt efflux from guard cells.

Patch-clamp experiments were performed on guard-cell protoplasts isolated from *Vicia faba* by a gentle procedure (no agitation) at the temperature of plant cultivation⁵. High-resistance membrane seals⁴ were formed between the patch pipette and the plasma membrane (Fig. 1). Fluxes of K⁺, Cl⁻ and malate²⁻ were recorded at the cellular and single-channel level using the whole-cell configuration as well as isolated membrane patches⁴.

Whole-cell currents were measured in the presence of 150 mM KCl in the pipette and 30 mM KCl in the bath (Fig. 2a, filled circles). The membrane conductance was studied as a function of the membrane voltage applied to the cytoplasmic side of the plasma membrane. The relationship between current and voltage depended on the voltage range considered. At membrane potentials below -100 mV, large inward currents occurred. They were identified as inward K⁺ currents by reversal potential analysis². Membrane potentials above -20 mV, particularly positive potentials, caused outward K⁺ currents; at $+50$ mV they were as large as 100 pA. In the voltage range between -100 mV and -80 mV, which corresponds to the resting potentials of guard cells⁶, the current was close to zero and the resistance of the plasma membrane was larger than 10 GΩ. Increasing the potential to values positive to -80 mV led to an inward current with a maximum conductance at -40 mV, the reversal potential for K⁺.

To test whether this current was carried by anions, K⁺ was replaced by Cs⁺ or N-methylglucamine (NMG), which are known to permeate K⁺ channels at negligible rates⁷ (Fig. 2a, open circles). Below -100 mV and above -20 mV, currents were much smaller than they were in the presence of K⁺, or even zero. Their fluctuations reached a maximum amplitude at -41 mV (Fig. 2b) and they disappeared at $+39$ mV (Fig. 2b), close to the Nernst potential for Cl⁻ ($+40$ mV). These results provide evidence that anions are the main charge carrier⁵ of the whole-cell currents in the presence of CsCl solutions. At -81 mV, the opening and closing of single anion channels could be clearly discerned, even in whole-cell recordings.

Received 23 June; accepted 16 August 1989.

- McMichael, A. J. et al. *Eur. J. Immun.* **9**, 205-210 (1979).
- Fithian, E. et al. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2541-2544 (1981).
- Small, T. N. et al. *J. Immun.* **138**, 2864-2868 (1987).
- Amiot, M. et al. *J. Immun.* **136**, 1752-1758 (1986).
- Calabi, F. & Milstein, C. *Nature* **323**, 540-543 (1986).
- Balk, S. P., Bleicher, P. A. & Terhorst, C. *Proc. natn. Acad. Sci. U.S.A.* **86**, 252-256 (1989).
- Terhorst, C. et al. *Cell* **23**, 771-780 (1981).
- Brenner, M. B. et al. *Nature* **325**, 689-694 (1987).
- Jitsukawa, S., Faure, F., Lipinski, M., Triebel, F. & Hercend, T. *J. exp. Med.* **166**, 1192-1197 (1987).
- Sieckman, P. P. et al. *Nature* **328**, 351-353 (1987).
- van de Rijn, M. et al. *Science* **226**, 1083-1085 (1984).
- Santos-Aguado, J., Barbosa, J. A., Biro, P. A. & Strominger, J. L. *J. Immun.* **141**, 2811-2818 (1988).
- van Agthoven, A. & Terhorst, C. *J. Immun.* **128**, 426-432 (1982).
- Strominger, J. L. *Cell* **57**, 895-898 (1989).
- Shivakumar, S., Tsokos, G. C. & Datta, S. K. *J. Immun.* **143**, 103-112 (1989).
- Miescher, G. C., Budd, R. C., Lees, R. K. & MacDonald, H. R. *J. Immun.* **138**, 1959-1967 (1987).
- Brenner, M. B., Strominger, J. L. & Krangel, M. S. *Adv. Immun.* **43**, 133-192 (1988).
- Bluestone, J. A., Cron, R. Q., Cotterman, M., Houlden, B. A. & Matis, L. A. *J. exp. Med.* **168**, 1899-1916 (1988).
- Janeway, C. A. Jr, Jones, B. & Hayday, A. C. *Immun. Today* **6**, 73-76 (1988).
- Lerch, P. G., van de Rijn, M., Schrier, P. & Terhorst, C. *Hum. Immun.* **6**, 13-30 (1983).
- Reinherz, E. R., Kung, P. C., Goldstein, G., Levey, R. H. & Schlossman, S. F. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1588-1592 (1980).
- Knowles, R. K. & Bodmer, W. F. *Eur. J. Immun.* **12**, 676-681 (1982).
- Martin, L. H., Calabi, F., Lefebvre, F.-A., Bilisand, C. A. G. & Milstein, C. *Proc. natn. Acad. Sci. U.S.A.* **84**, 9189-9193 (1987).
- Brody, F. M. & Parham, P. *J. Immun.* **128**, 129-135 (1982).
- Spits, H. et al. *Hybridoma* **2**, 423-437 (1983).
- Band, H. et al. *Science* **238**, 682-694 (1987).
- Lanier, L. L., Ruitenberg, J. J., Allison, J. P. & Weiss, A. in *Leukocyte Typing III* (ed. McMichael, A. J.) 175-178 (Oxford University Press, 1987).
- Arruffo, A. & Seed, B. *J. Immun.* **143**, 1723-1730 (1989).
- de Waal Malefyt, R. et al. *J. Immun.* **142**, 3634-3642 (1989).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, New York, 1982).
- Potter, H., Weir, L. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7161-7165 (1984).
- Wigler, M. et al. *Cell* **16**, 777-784 (1979).
- Moingeon, P. et al. *Nature* **339**, 312-314 (1989).

ACKNOWLEDGEMENTS. We thank Drs A. Arruffo and B. Seed for the CD1a, b and c cDNA clones, Dr Ken-ichi Arai for the pSRα-NEO vector, Dr P. H. Schur for specimens of human peripheral blood, Dr H. Band for reviewing this manuscript, and for their gifts of monoclonal antibodies, Drs W. Knapp (1OC3 and 7C6), M. Amiot and L. Boumsell (L161), R. Knowles (M241), B. Seed (NU-T2), R. Kurrie (BMA031), T. Hercend (anti-TiV), and J. De Vries (SPV-T3b). Technical assistance was provided by A. Reinech and S. Pilsarczyk. This work was supported by the NIH and ACS.

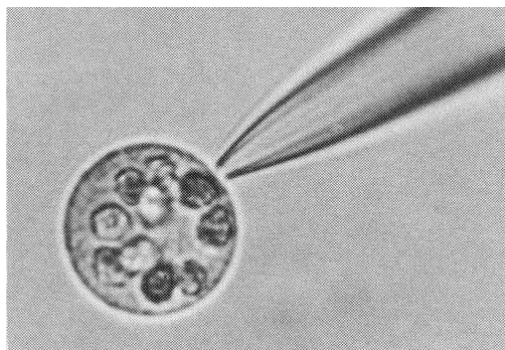


FIG. 1 Micrograph (Nomarski interference contrast) of a guard-cell protoplast from *Vicia faba*.

METHODS. Guard-cell protoplasts were isolated without agitation by overnight incubation at 20 °C in small Petri dishes containing 2.5 ml of 400 mM D-mannitol, 10 mM Na-ascorbate, 1 mM CaCl_2 , 2% (w/v) cellulase (Onozuka RS), 1% pectinase (for example, Mazerozyme), 0.5% BSA, 0.05% kanamycin sulphate, pH 5.5, with osmolality adjusted to produce incipient plasmolysis of guard cells. Protoplasts were purified by filtration and centrifugation⁵.

In inflated guard cells, malate is the major anion. Replacing chloride by malate left the voltage dependence of the anion current essentially unchanged. The maximum current still occurred close to -40 mV, but was reduced to values below 10 pA. After raising concentrations of cytosolic free Ca^{2+} to $>5 \mu\text{M}$, anion currents could be measured at voltages negative to -60 mV. At a Ca^{2+} concentration in the pipette of $>2 \text{ mM}$, whole-cell anion currents did not exceed 30 pA in the presence of CsCl solutions.

After withdrawal of the pipette from the membrane surface,

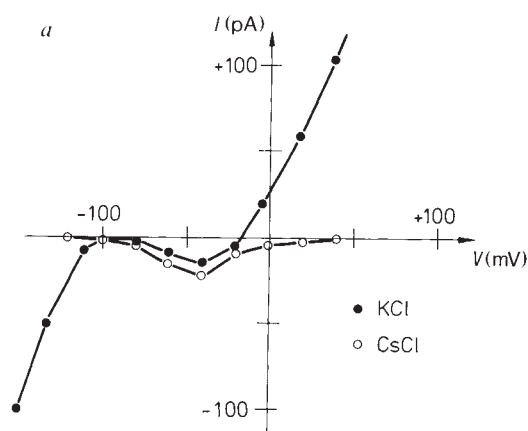
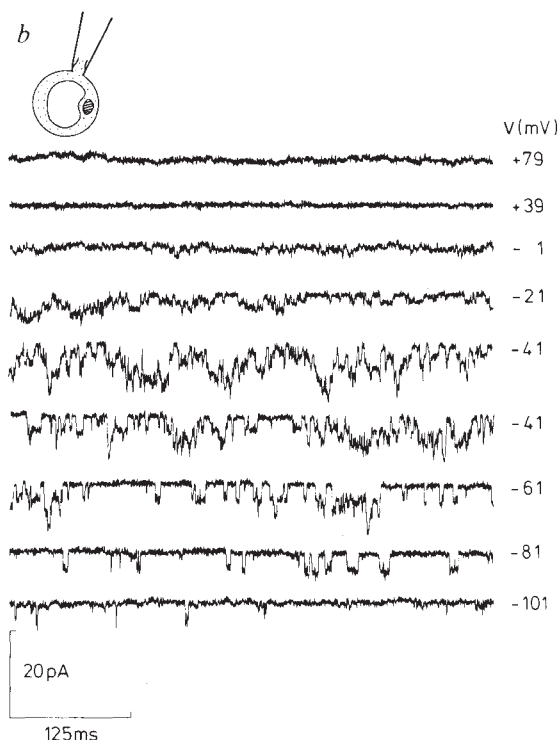


FIG. 2 Whole-cell currents of guard-cell protoplasts. *a*, ●, current-voltage (I - V) relationship of protoplasts bathed in 30 mM KCl, 1 mM CaCl_2 , 2 mM MgCl_2 , 10 mM HEPES-Tris buffer, pH 5.5. Pipette solution, 150 mM KCl, 2 mM MgCl_2 , 2 mM Mg ATP, 1 mM EGTA, 10 mM HEPES-Tris buffer, pH 7.2. Current-voltage relationships were determined by continuously increasing the voltage from -150 mV to +50 mV. Voltage and current were plotted on an x-y screen and corrected for the (ohmic) resistance of the membrane seal. Downward currents corresponded to cation influx into (or anion efflux from) the protoplast or both. ○, current-voltage relationship of protoplasts bathed in 30 mM CsCl, 1 mM CaCl_2 , 2 mM Mg MgCl_2 , 10 mM HEPES-Tris buffer, pH 5.6. Pipette solution was 150 mM CsCl, 2 mM MgCl_2 , 2 mM Mg-ATP, 1 mM EGTA, 10 mM HEPES-Tris buffer, pH 7.2. *b*, Whole-cell currents recorded at various voltages V in asymmetrical CsCl solutions corresponding to those in Fig. 2*a* (○). Membrane potentials were controlled by an EPC-7 patch-clamp amplifier (List Electronics, Darmstadt), and data were stored in digital form on a video-tape recorder. Off-line analysis was performed on a PDP-11/73 computer after filtering the data at 2 kHz (8-pole Bessel filter). All membrane potentials refer to the cytoplasmic side of the plasma membrane. Experiments were performed at room temperature (20 ± 2 °C).

single anion channels and K^+ channels were recorded in the outside-out patch configuration in K^+ salt solutions (Fig. 3). To avoid interferences by the activity of K^+ channels, we limited the investigation to voltages positive to -90 mV, at which inward K^+ channels were closed. Between -90 mV and -40 mV, the membrane current was dominated by the opening of single anion channels. Downward current deflections represented single-channel currents of 3.9 pA at -90 mV. In excised patches, anion channels were most active when the concentration of free Ca^{2+} on the cytoplasmic side of the membrane was raised to 3-5 μM . Channel activity occurred in bursts separated by silent periods of 0.2 to 0.9 s duration. At voltages positive to -20 mV, the opening of outward K^+ channels appeared as upward current deflections. Multiple current levels indicated the simultaneous opening of several K^+ channels in the patch. At voltages above +20 mV, currents through the plasma membrane were dominated by the activity of K^+ channels, with a current amplitude of 2 pA at +70 mV (Fig. 3*a*).

For single anion channels, a main conductance level of 39 pS in 100 mM KCl was calculated from current-voltage measurements (Fig. 3*b*). Occasionally, channel openings at lower conductance levels could be observed, as shown in Fig. 3*c*. To assess the selectivity of the anion channels, Cl^- was replaced by several other anions. In reversal-potential measurements we found a permeability sequence of $\text{NO}_3^- > \text{Cl}^- > \text{malate}^{2-}$. Anion permeation was blocked by the addition of 10 μM Zn^{2+} , a well-known blocker of anion channels⁸ (Fig. 4).

The anion currents we observed differ significantly from the anion currents recently described in the plasma membrane of plant cells. Anion currents in suspension-cultured cells of *Asclepias tuberosum* were elicited by hyperpolarizing instead of depolarizing voltages and had a magnitude of 300-400 pA in the whole-cell configuration⁹. The corresponding single-channel conductance was 100 pS compared with 39 pS for guard cells,



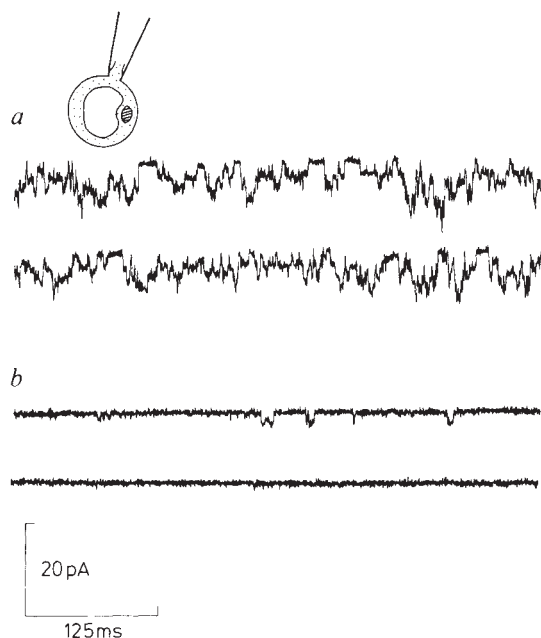
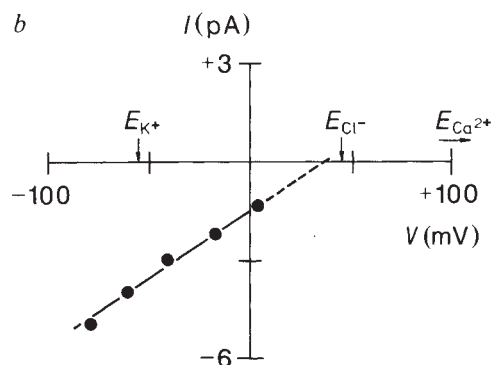
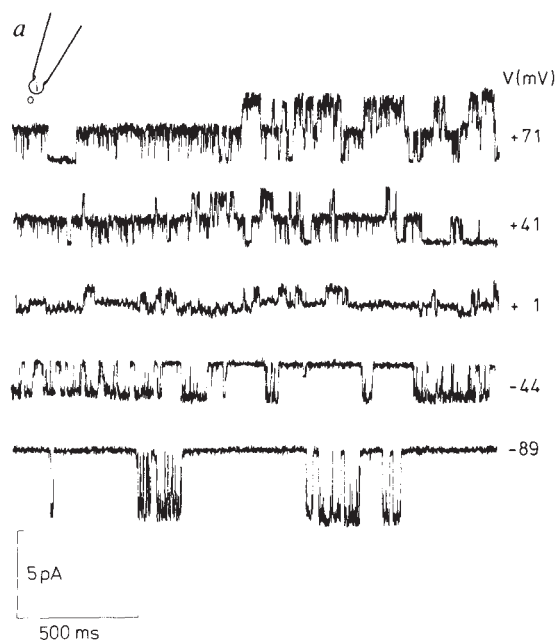


FIG. 4 Block of anion channels by extracellular Zn^{2+} . *a*, Currents through several anion channels were recorded in the whole-cell configuration. Protoplasts were bathed in 30 mM CsCl, 1 mM CaCl_2 , 2 mM MgCl_2 , 10 mM HEPES-Tris buffer, pH 5.6. Pipette solution, 150 mM CsCl, 2 mM MgCl_2 , 2 mM MgATP, 1 mM EGTA, 10 mM HEPES-Tris buffer, pH 7.2. The membrane potential was held at -41 mV. *b*, Addition of $10 \mu\text{M}$ ZnCl_2 to the bath solution blocked the activity of anion channels.

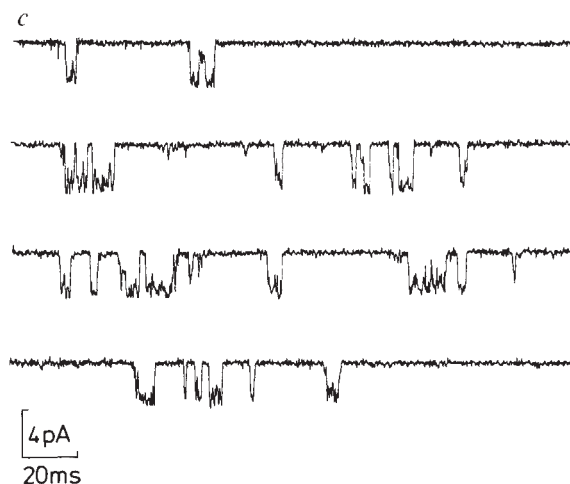


FIG. 3 Single-channel recordings from excised patches. *a*, Single K^+ - and anion-selective channels measured in outside-out patches in an asymmetrical solution of KCl. Pipette solution, 100 mM KCl, 2 mM MgCl_2 , 2 mM Mg-ATP, 1 mM EGTA, 10 mM HEPES-Tris buffer, pH 7.2; Bath solution, 10 mM KCl, 1 mM CaCl_2 , 2 mM MgCl_2 , 10 mM HEPES-Tris buffer, pH 5.8. Upward current deflections at positive membrane voltages correspond to K^+ currents flowing through open K^+ channels. Current recordings at negative voltages are dominated by the opening and closing of single anion channels. Note the gating of both K^+ - and anion channels at $V = +1$ mV. *b*, Single-channel current-voltage relationship. Single-channel current-voltage curve observed in a 10-fold gradient of KCl. Pipette solution, 100 mM KCl, 2 mM MgCl_2 , 2 mM Mg-ATP, 0.1 mM CaCl_2 , 1 mM EGTA, 10 mM HEPES-Tris buffer, pH 7.3; Bath solution, 10 mM KCl, 1 mM CaCl_2 , 2 mM MgCl_2 , 10 mM HEPES-Tris buffer, pH 5.5. The Nernst potentials for Cl^- , K^+ and Ca^{2+} were $E_{\text{Cl}} = -44.1$ mV, $E_{\text{K}^+} = -55.0$ mV and $E_{\text{Ca}^{2+}} > +100$ mV, respectively (values corrected for ionic activities). Slope conductance, 39 pS. *c*, Currents through single anion-channels. Single-channel currents were recorded at -82 mV in the presence of the non-permeating cation NMG. Pipette solution, 100 mM NMG-Cl, 2 mM MgCl_2 , 2 mM Mg-ATP, 1 mM EGTA, 10 mM HEPES-Tris buffer, pH 7.4. Bath solution, 10 mM NMG-Cl, 1 mM CaCl_2 , 2 mM MgCl_2 , 10 mM HEPES-Tris buffer, pH 5.5. Note the occurrence of channel substates in the first record from top.

indicating a striking discrepancy between anion transport in non-differentiated suspension-cultured cells and guard-cell protoplasts from developed leaves.

Our results also disagree with those of Schroeder and Hagiwara¹⁰, who showed instantaneously activated, almost voltage-insensitive anion conductances in guard-cell protoplasts from *Vicia faba*. These anion currents were observed at high concentrations of cytoplasmic Ca^{2+} and exceeded outward K^+ currents by a factor of three (0.7 nA at -70 mV). Outward anion currents of that magnitude did not occur in guard cells embedded in the epidermis, where anion release during stomatal closure was equivalent to 30 – 40 pA¹¹. Therefore, the anion channels shown here provide suitable molecular pathways to determine the magnitude of anion release from guard cells. Whether Ca^{2+} -activated anion conductances function in addition or alternatively to the voltage-regulated anion channels described here needs to be investigated further.

An early proposal was that the release of salts by guard cells is determined primarily by the anion permeability of the plasma membrane and that an increase in this permeability was sufficient to initiate stomatal closure^{3,12}. The anion channels that we have described here would allow the depolarization required for salt release from guard cells. The ensemble conductance of the anion channels permeable to Cl^- and malate²⁻ can account for the observed rates of anion loss during stomatal closure¹¹. The anion channels in guard cells indeed seem to provide a mechanism for controlled salt release. We propose that this

type of anion channel also functions in other types of plant cells, where it may be involved in the release of salts for the purpose of osmoregulation and for the control of turgor or cell volume. □

Received 14 June; accepted 21 August 1989.

- Schroeder, J. I., Hedrich, R. & Fernandez, J. M. *Nature* **312**, 361–362 (1984).
- Schroeder, J. I., Raschke, K. & Neher, E. *Proc. nat. Acad. Sci. U.S.A.* **84**, 4108–4112 (1987).
- Raschke, K. *Encyclopedia of Plant Physiology Vol. 7. Physiology of Movements* **383–441** (Springer, Berlin/Heidelberg/New York, 1979).
- Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. *Pflügers Arch. ges. Physiol.* **391**, 85–100 (1981).
- Raschke, K. & Hedrich, R. *Meth. Enzym.* Vol. 174 (in the press).
- Blatt, M. R. *Planta* **170**, 272–287 (1987).
- Schroeder, J. I. *J. gen. Physiol.* **92**, 667–683 (1988).
- Hille, B. *Ionic Channels of Excitable Membranes*. (Sinauer, Sunderland, Massachusetts, 1984).
- Schäuf, C. L. & Wilson, K. J. *Plant Physiol.* **85**, 413–418 (1987).
- Schroeder, J. I. & Hagiwara, S. *Nature* **338**, 427–430 (1989).
- Van Kirk, C. A. & Raschke, K. *Plant Physiol.* **61**, 474–475 (1978).
- MacRobbie, E. A. C. *Bot. Acta* **101**, 140–148 (1988).

ACKNOWLEDGEMENTS. We thank I. Baumann for technical assistance, Dr W. Stühmer for comments on the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft and a fellowship from the Stiftung Volkswagenwerk granted to K.R., who also thanks Dr E. Neher, Max-Planck-Institut für biophysikalische Chemie, Göttingen, for hospitality in his laboratory.

Trans-dominant inactivation of HTLV-I and HIV-1 gene expression by mutation of the HTLV-I Rex transactivator

Laurence Rimsky, Madeleine Duc Dodon*, Eric P. Dixon & Warner C. Greene

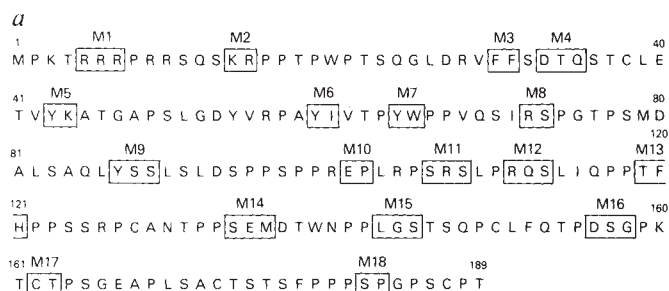
Howard Hughes Medical Institute, Departments of Medicine and Microbiology-Immunology, Duke University Medical Center, Durham, North Carolina 27710, USA

* Immuno-Virologie Moléculaire et Cellulaire, UMR30-CNRS/UCBL, Faculté de Médecine A Carrel, 69372 Lyon Cedex 8, France

THE *rex* gene of the type I human T-cell leukaemia virus (HTLV-I) encodes a phosphorylated nuclear protein of relative molecular mass 27,000 which is required for viral replication^{1–3}. The Rex protein acts by promoting the cytoplasmic expression of the incompletely spliced viral messenger RNAs that encode the virion structural proteins^{4,5}. To identify the biologically important peptide domains within Rex, we introduced a series of mutations throughout its sequence. Two distinct classes of mutations lacking Rex biological activity were identified. One class corresponds to *trans*-dominant repressors as they inhibit the function of the wild-type Rex protein. The second class of mutants, in contrast, are recessive negative, rather than dominant negative, as they are not appropriately targeted to the cell nucleus. These results indicate the presence of at least two functionally distinct domains within the Rex protein, one involved in protein localization and a second involved in effector function. The *trans*-dominant Rex mutants may represent a promising new class of anti-viral agents.

Oligonucleotide-directed mutagenesis was employed to alter the primary sequence of the *rex* gene at 18 discrete sites (Fig. 1a). The boxed amino acids were replaced with a dipeptide, aspartic acid-leucine, by insertion of an in-frame oligonucleotide duplex that contains a diagnostic *Bgl*III restriction site. Each of these *rex* mutants were then inserted into the pBC12/CMV eucaryotic expression vector⁶ and the mutations verified by DNA sequencing. For functional analysis, a Rex-responsive reporter plasmid, pgTAX-LTR, was prepared (Fig. 1b). The pgTAX-LTR vector contains the two protein-coding exons of the *tax* gene separated by the HTLV-I *env* gene and the complete HTLV-I 3' long terminal repeat (LTR). Contained within this 3' LTR is a *cis*-

acting Rex response element (RexRE) which seems to correspond to a complex RNA stem-loop structure^{5,7}. Either direct or indirect binding of Rex to this RexRE probably underlies its post-transcriptional mechanism of action. The pgTAX-LTR vector does not itself produce Rex, because of a mutation introduced at the Rex initiator codon. In the absence of Rex, pgTAX-LTR produces the Tax protein, reflecting the translation of the spliced HTLV-I mRNA species that lacks virtually all of the *env* sequences. When pgTAX-LTR is co-transfected with the *rex* expression plasmid, pREX, however, synthesis of the HTLV-I Env protein is activated. This protein (relative molecular mass 68,000 (68K)) is readily identified by immunoprecipitation with the anti-HTLV-I envelope monoclonal antibody, 0.5- α (ref. 8) (Fig. 2a; lane 1). In contrast, no HTLV-I Env protein is detected when pREX is replaced by either pREV (lane 2) or pCMV-IL-2 (lane 3), which encode the HIV-1 Rev protein and human interleukin-2 polypeptide respectively. Also, no Env protein is identified when cells are transfected with pREX in the absence of pgTAX-LTR (lane 4). HTLV-I Env expression by pgTAX-LTR, therefore, is specifically induced in the presence of the wild-type Rex protein.



b

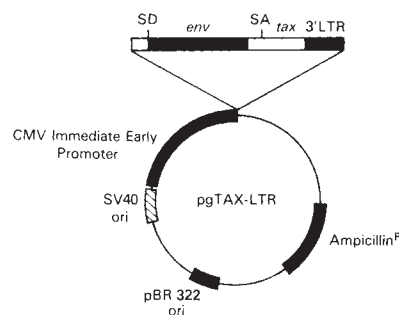


FIG. 1 a, Amino-acid sequence of the HTLV-I Rex protein and location of the 18 point mutations introduced. Using oligonucleotide-directed mutagenesis in the M13 bacteriophage²¹, nucleotides encoding each of the boxed amino acids were removed and replaced in frame by an oligonucleotide encoding the aspartic acid-leucine dipeptide. Each mutant was subsequently inserted into the pBC12/CMV expression vector and the presence of the mutation confirmed by DNA sequencing of the double-stranded supercoiled templates. b, Structure of the Rex-responsive pgTAX-LTR expression vector. The 3' end of the CR-1 HTLV-I provirus (a gift from Dr Flossie Wong-Staal) from the HindIII site at map position 5,013 (ref. 22) through to the 3' LTR, was inserted into the pBC12/CMV expression vector. This fragment contains the two coding exons for Tax (white boxes), the complete *env* gene, and the entire 3' LTR, including the Rex response element⁵. Although this vector contains the coding region for Rex, it does not synthesize this protein because of the introduction of a mutation at the *Sph*I site, which is coincident with the Rex translation initiation codon¹⁴. Expression of these HTLV-I sequences is promoted by the immediate early region of the human cytomegalovirus (CMV) and additional polyadenylation sequences provided by the 3' region of the rat pre-proinsulin gene⁶. This vector produces Env only in the presence of Rex, but Tax is synthesized in the presence or absence of Rex.

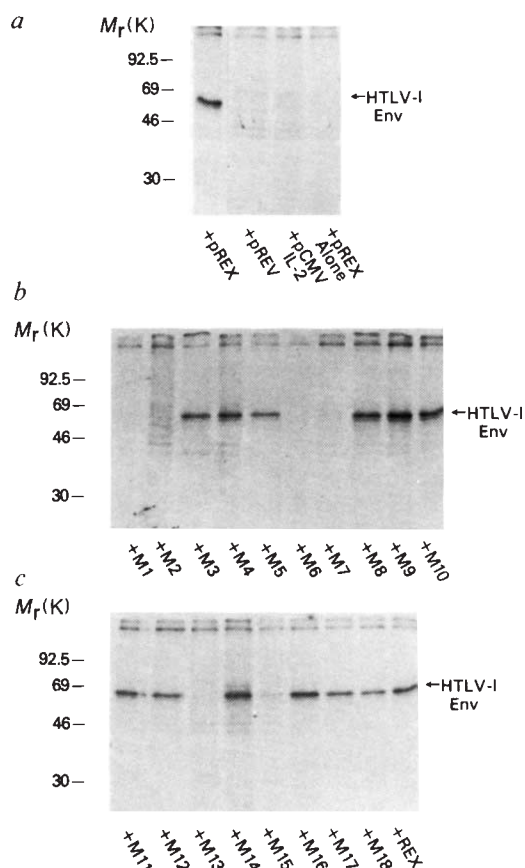


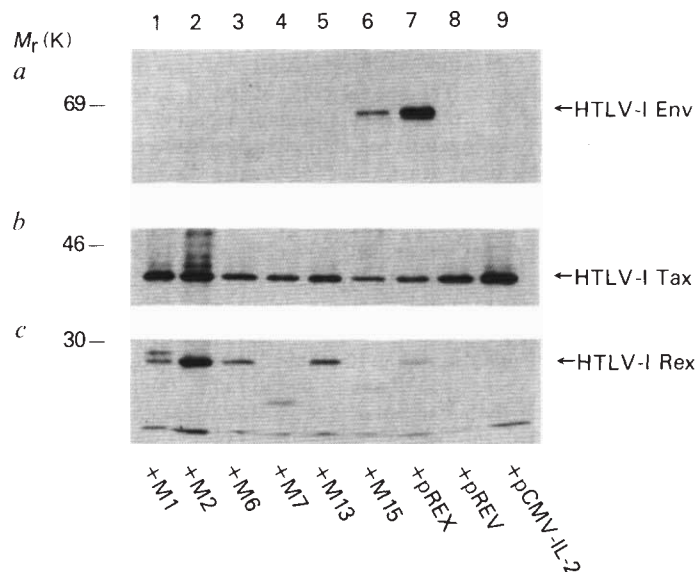
FIG. 2 *a*, Rex, but not Rev or IL-2, activates the expression of HTLV-I Env protein by the pgTAX-LTR vector. Subconfluent cultures of COS cells were co-transfected with pgTAX-LTR and pREX (ref. 14), pREV (ref. 15), or pCMV-IL-2 (ref. 6) using DEAE-dextran²³. All plasmids were added at a concentration of $1.25 \mu\text{g ml}^{-1}$. In the final lane, cells were transfected with pREX in the absence of pgTAX-LTR. Forty-eight hours after transfection, cells were metabolically labelled with [³⁵S] cysteine for 2 hours, subcellular extracts were fractionated. Samples were immunoprecipitated with 0.5- α human monoclonal antibody, which specifically reacts with the HTLV-I envelope protein⁸. Immunoprecipitates were analysed on SDS-10% polyacrylamide gels. The migration of molecular weight standards is indicated on the left. *b*, Analysis of biological activity of the *rex* mutants. After insertion into the pBC12/CMV expression vector, each of the *rex* mutants (designated M1–M18, see Fig. 1a) was co-transfected with pgTAX-LTR and the COS cell cultures were analysed for Rex-dependent HTLV-I Env protein expression as described above.

Each of the *rex* mutations was tested for biological activity by co-transfection with pgTAX-LTR vector. Twelve of these *rex* mutants displayed a wild-type phenotype. In contrast, five mutants (M1, M2, M6, M7 and M13) lacked apparent biological activity and one mutant (M15) was only partially functional (Fig. 2b). COS cells were next co-transfected with these six defective *rex* mutants and the pgTAX-LTR vector followed by the simultaneous analysis of HTLV-I Env, Tax and Rex protein expression (Fig. 3a–c). Although HTLV-I Env was only detected in the presence of the wild-type Rex protein (Fig. 3a; lane 7) or the partially active M15 mutant (Fig. 3a; lane 6), the 40K Tax protein was detected in all of the cultures (Fig. 3b). These findings indicate that the lack of *env* gene expression observed with the M1, M2, M6, M7 and M13 mutants is due to the specific loss of Rex biological activity, rather than to non-specific, toxic effects of these proteins in the transfected COS cell cultures. Each of the mutant Rex proteins was also directly identified by immunoprecipitation, indicating that each was expressed in a stable manner (Fig. 3c). The M2, M6, and M13 Rex mutants migrated in a manner indistinguishable from the wild-type Rex protein (Fig. 3c; lanes 2, 3, 5 and 7). Interestingly, the M7 and M15 proteins had a smaller apparent molecular weight (lanes 4 and 6), and the M1 mutant protein travelled as an electrophoretic doublet (lane 1). Sequencing of the entire protein-coding region in the M1, M7 and M15 mutants failed to reveal any changes other than the specific mutations introduced. These apparent differences in size, therefore, probably reflect altered post-translational processing of the mutant Rex proteins.

The *in situ* immunofluorescent staining of cells transfected with the wild-type *rex* or the biologically inactive M6, M7 and M13 Rex mutants revealed expression of these proteins in the nucleoli and nuclei of the transfected cells (Fig. 4). In sharp contrast, the M1 Rex mutant was detectable only in the cytoplasm, but the M2 Rex mutant was distributed homogeneously throughout the cell. In accordance with these results, the M1 and M2 mutations altered basic amino-acid residues located within a positively charged peptide domain that has been previously shown to function as a nucleolar localization signal³. Interestingly, the partially active M15 mutation led to a nuclear, but nucleolar-excluded, pattern of expression (Fig. 4). These findings raise the possibility that residues apart from the basic amino acids at the N-terminus may be involved in the nucleolar expression of Rex.

The Rex mutants were next tested for their capacity to block the biological action of the wild-type HTLV-I Rex and HIV-1

FIG. 3 Simultaneous analysis of HTLV-I Env, Tax, and Rex protein production in COS cells co-transfected with pgTAX-LTR ($1.25 \mu\text{g ml}^{-1}$) and the vectors encoding the biologically inactive (M1, M2, M6, M7, M13) or impaired (M15) mutant Rex proteins ($1 \mu\text{g ml}^{-1}$) or the wild-type pREX, pREV, and pCMV-IL-2 vectors ($1 \mu\text{g ml}^{-1}$). *a*, Analysis of HTLV-I Env production using the 0.5- α antibody for immunoprecipitation. *b*, Analysis of Tax protein production by pgTAX-LTR, which occurs in the absence or presence of Rex, using anti-Tax peptide specific antisera. *c*, Immunoprecipitation analysis of wild-type and mutant Rex protein production. In each case, the same radioactively labelled cellular extract was used for the three immunoprecipitations; only the relevant region of each of the resultant autoradiographs is presented. Immunoprecipitations were analysed on SDS-10% polyacrylamide gels.



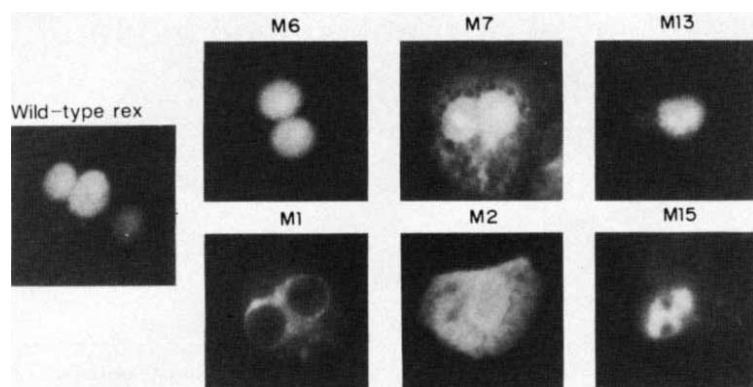


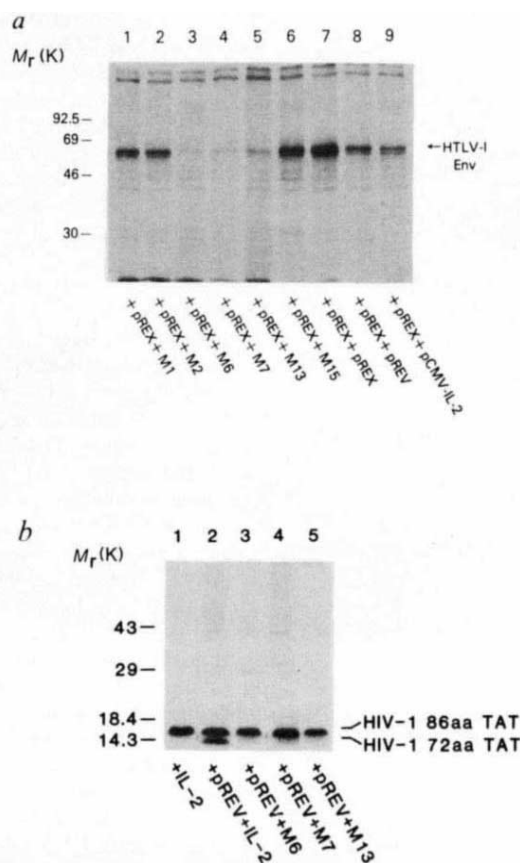
FIG. 4 Subcellular localization of the HTLV-I Rex mutants by immunofluorescence. COS cells were transfected with the expression plasmids indicated and fixed with paraformaldehyde 48 h later. The cells were then sequentially stained with rabbit anti-Rex peptide antiserum (1:100 dilution) and goat anti-rabbit IgG conjugated to rhodamine²³.

Rev proteins mediated through the RexRE and RevRE⁹⁻¹³ (Fig. 5). When co-transfected with pgTAX-LTR and pREX in COS cells (panel a), a 10-fold molar excess of M6, M7 and M13 mutants displayed a dominant negative phenotype, as the action of the wild-type Rex protein was markedly inhibited (lanes 3-5). In contrast, the M1, M2 and M15 proteins functioned as recessive negative mutants, because the action of the wild-type Rex protein was not altered (lanes 1, 2 and 6). Similarly, the Rev protein of HIV-1 did not interfere with the action of the HTLV-I Rex protein (lane 8), nor did IL-2 (lane 9). We have previously shown that the HTLV-I Rex protein can functionally replace the HIV-1 Rev protein¹⁴, so we next investigated whether *trans*-dominant Rex mutants could block the function of Rev in the

HIV-1 system (Fig. 5b). As previously described, for cells transfected with the pgTAT vector¹⁵, Rev induces the expression of a truncated (72 amino acids) rather than full-length (86 amino acids) form of the HIV-1 Tat protein (lane 2). When co-transfected with pgTAT and pREV (ref. 15) in COS cells, a 10-fold molar excess of M6, M7 or M13 Rex mutants inhibited the action of the Rev protein (lane 3, 4 and 5), as shown by the diminished expression of the 72-amino-acid form of the Tat protein.

The ability of these *trans*-dominant Rex mutants to block HIV-1 viral replication was also studied (Fig. 5c). Replication of a Rev-deficient HIV-1 provirus, pHXB2-Bam-p3 (ref. 14), in the presence of Rex and graded amounts of *trans*-dominant Rex

FIG. 5 a, Analysis of the ability of the rex mutants to inhibit function of the wild-type Rex protein. COS cultures were co-transfected with three plasmids, including pgTAX-LTR (1.25 $\mu\text{g ml}^{-1}$), pREX (0.1 $\mu\text{g ml}^{-1}$) and 1 $\mu\text{g ml}^{-1}$ of the Rex mutants (lane 1-6), pREX (lane 7), pREV (lane 8) or pCMV-IL-2 (lane 9). HTLV-I Env production was analysed by immunoprecipitation with the 0.5- α monoclonal antibody and then analysed by electrophoresis through an SDS-10% polyacrylamide gel. b, The HTLV-I Rex trans-dominant mutants inhibit the function of the HIV-1 Rev protein. COS cells were co-transfected with pgTAT, pREV, and a 10-fold molar excess of pBC/CMV IL-2 or the M6, M7 and M13 *trans*-dominant Rex mutants. After 48 h in culture and biosynthetic labelling with [³⁵S]cysteine, cellular extracts were assayed by immunoprecipitation for Rev-induced production of the truncated 72-amino-acid form of the Tat protein (72aa, lane 2; ref. 15). At a 10:1 molar ratio, the M6, M7 and M13 (lanes 3-5) completely inhibited the action of the HIV-1 Rev protein, because only the full-length 86-amino-acid (86aa) form of the Tat protein was detected. c, The HTLV-I *trans*-dominant Rex mutants block Rex rescue of the replication of a Rev-deficient HIV-1 provirus. COS-cells were co-transfected with the *rev*-deficient HIV-1 proviral plasmid, pHXB2-Bam-p3 (ref. 24) and pREX in the presence of the indicated fold-excess of the M1, M6, M7 and M13 mutants. Total DNA concentration in the transfection cocktail was maintained at a constant level by the addition of varying amounts of the pBC/CMV-IL-2 parental vector. Three days after transfection, supernatant levels of the



HIV-1 p24 Gag protein were measured by ELISA (Coulter Immunology kit). The M6, M7 and M13 mutants produced dose-related inhibition of HIV-1 p24 production, but the recessive negative M1 mutant did not.

mutants was studied by transfection of COS cells with these plasmids. As indicated by the supernatant levels of synthesized HIV-1 p24 Gag protein, the *trans*-dominant Rex mutants (M6, M7 and M13) produced dose-related inhibition of HIV-1 replication. In contrast, the recessive negative M1 mutant of Rex was without significant effect.

These findings, together with those from the various *rex* mutants, suggest that there are at least two different functional peptide domains within the Rex protein. One domain defined by the M1 and M2 recessive negative mutations is located at the N-terminus, involves amino acids 5–7 and 14–15, and seems to play a part in targeting of the protein to the nucleus. It is also possible that this positively charged domain may be involved in Rex binding, either directly to the RexRE or to other proteins that directly contact this RNA element. A second functional domain encompasses amino acids 59–60 (tyrosine-isoleucine), 64–65 (tyrosine-tryptophan) and 119–121 (threonine-phenylalanine-histidine). Alterations at each of these discrete sites (M6, M7 and M13 mutants) lead to the production of Rex proteins that lack biological activity and also display *trans*-dominant repressor properties. Of note is the fact that the five different mutations that retain full biological activity separate M6/M7 and M13 in the linear sequence of the Rex protein, raising the possibility that the interaction of these discrete regions requires proper protein folding. This domain seems to be important in the effector function of the Rex protein, which regulates HTLV-I structural gene expression.

These *trans*-dominant repressors of the HTLV-I *rex* gene thus join a small but expanding group of dominant negative mutant proteins which represent a novel class of theoretical anti-viral reagents. Specifically, *trans*-dominant mutants of the VP-16 protein of the herpes simplex virus¹⁶, the Tax protein of HTLV-II (ref. 17), the Rev protein of HIV-1 (ref. 18), and the Tat protein of HIV-1 (ref. 19) have now been described. Should effective and safe delivery systems be developed, each of these mutant genes or proteins could in theory be used as 'intracellular immunogens'²⁰ to block the replication of the respective virus by inhibiting the essential function of its wild-type counterpart protein. Certain *trans*-dominant proteins might also prove active against more than one type of virus. Specifically, we have demonstrated that the HTLV-I Rex *trans*-dominants are active as repressors of Rev function in the HIV-1 system. These dominant negative Rex proteins, therefore, seem to be capable of mediating anti-viral effects in both the HTLV-I and HIV-1 systems. □

Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues

Gunnar von Heijne

Department of Molecular Biology, Karolinska Institute, Centre for Biotechnology, NOVUM, S-141 52 Huddinge, Sweden

POSITIVELY charged amino acids have been shown to be important elements in targeting-peptides that direct proteins into mitochondria, nuclei, and the secretory pathways of both prokaryotic and eukaryotic cells¹. The 'positive-inside' rule, which observes that regions of polytopic (multi-spanning) membrane proteins facing the cytoplasm are generally enriched in arginyl and lysyl residues whereas translocated regions are largely devoid of these residues^{2,3}, implies that the distribution of positively charged amino acids may also be a major determinant of the transmembrane topology of integral membrane proteins. If this is indeed the case, it should be possible to predictably alter the topology of a polytopic protein by site-directed insertions and/or deletions of positively charged residues in critical locations. I now describe a derivative of *Escherichia coli* leader peptidase, a polytopic inner-membrane protein, that switches from *sec*-gene-dependent membrane insertion with a $N_{out}-C_{out}$ transmembrane topology to *sec*-gene-independent insertion with a $N_{in}-C_{in}$ topology in response to the addition of four positively charged lysines to its N terminus.

Leader peptidase (Lep) from *E. coli*, Fig. 1, has two apolar transmembrane segments (H1: amino acids 4–22, H2: amino acids 62–76), is made without a cleavable signal sequence, and is oriented with its N- and C-terminal domains facing the periplasm (refs 4, 5 and see below). Its membrane assembly has been extensively studied; both apolar regions H1 and H2 are required for assembly^{6–8} and like soluble periplasmic proteins, Lep depends on the products of the *secA* and *secY* genes for proper membrane integration⁹. To test the predictive power of the positive inside rule, I have sought to invert the topology of Lep by changing the distribution of positively charged residues from a typical $N_{out}-C_{out}$ pattern to a $N_{in}-C_{in}$ pattern by adding four positively charged lysines to the N terminus (between residues 4 and 5, called mutant 4K) and simultaneously removing the highly charged segment 30–52 from region P1 between H1 and H2 (mutant $\Delta 30-52$).

As shown previously⁸, the 4K and $\Delta 30-52$ mutations do not by themselves affect the topology of Lep as judged by a protease-accessibility assay. The double-mutant 4K/ $\Delta 30-52$, however, is completely resistant to proteases added to cells with permeabilized outer membranes (data not shown). This is consistent with an 'inverted' ($N_{in}-C_{in}$) topology where the large P2 domain is protected from protease attack by the inner membrane, but the location of region P1 remains undefined.

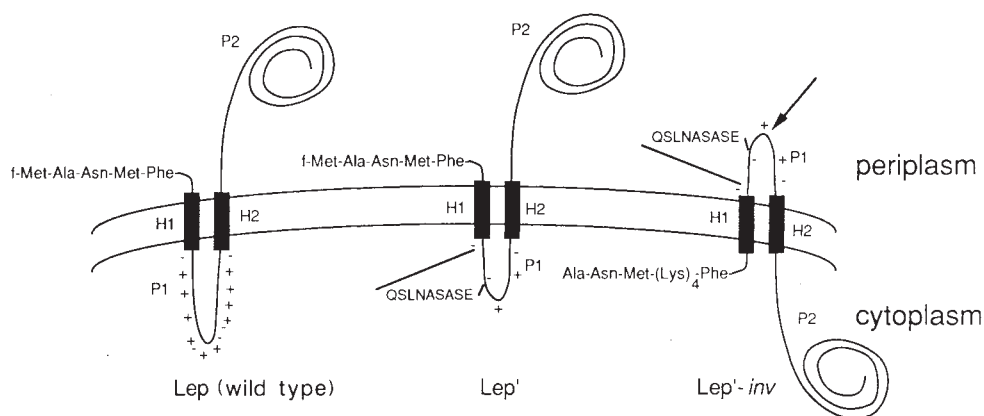
The P1 region might, even if translocated across the inner membrane, be too short and/or too tightly bound to the membrane surface to be protease sensitive. Mutant Lep' was created by adding to the $\Delta 30-52$ mutant a short spacer of eight mainly polar and one negatively charged amino acids (Gln-Ser-Leu-Asn-Ala-Ser-Ala-Ser-Glu, shown in single-letter code in Fig. 1) between residues 23 and 24 in P1. As shown in Fig. 2, Lep' behaves like its parent, the $\Delta 30-52$ mutant, and like wild-type Lep in the protease assay, demonstrating that the large C-terminal domain P2 is translocated across the inner membrane (and is therefore trypsin sensitive). When four positively charged lysines were added to the N-terminus of Lep', however, the resulting protein (Lep'-*inv*) was only trimmed to a slightly smaller form by trypsin as would be expected of a molecule with $N_{in}-C_{in}$ topology. The protected fragment has a molecular

Received 26 June; accepted 24 August 1989.

- Inoue, J., Yoshida, M. & Seiki, M. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3653–3657 (1987).
- Kiyokawa, T. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8359–8363 (1985).
- Siemi, H. et al. *Cell* **55**, 197–209 (1988).
- Hidaka, M., Inoue, J., Yoshida, M. & Seiki, M. *EMBO J.* **7**, 519–523 (1988).
- Hanly, S. et al. *Genes Dev.*, in the press.
- Cullen, B. R. *Cell* **46**, 973–982 (1986).
- Seiki, M., Inoue, J., Hidaka, M. & Yoshida, M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7124–7128 (1988).
- Matsushita, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2672–2676 (1986).
- Malim, M. H., Hauber, J., Le, S. Y., Maizel, J. V. & Cullen, B. R. *Nature* **338**, 254–257 (1989).
- Hadzopoulou-Cladaras, M. et al. *J. Virol.* **63**, 1265–1274 (1989).
- Fleber, B. K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T. & Pavlakis, G. N. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1495–1499 (1989).
- Gutman, D. & Goldenberg, C. J. *Science* **241**, 1492–1495 (1988).
- Emmerman, M., Vazeux, R. & Peden, K. *Cell* **57**, 1155–1165 (1989).
- Rimsky, L. et al. *Nature* **335**, 738–740 (1988).
- Malim, M. H., Hauber, J., Fenrick, R. & Cullen, B. R. *Nature* **335**, 181–183 (1988).
- Triesenberg, S. J., Kingbury, R. C. & McKnight, S. L. *Genes Dev.* **2**, 718–729 (1988).
- Wachsman, W. et al. *Science* **235**, 674–677 (1987).
- Malim, M., Böhnlein, S., Hauber, J. & Cullen, B. R. *Cell* **58**, 205–214 (1989).
- Green, M., Ishino, M. & Loerjens, P. M. *Cell* **58**, 215–223 (1989).
- Baltimore, D. *Nature* **335**, 395–396 (1988).
- Nakamaye, G. & Eckstein, F. *Nucleic Acids Res.* **14**, 9679–9698 (1986).
- Seiki, M., Hattori, S., Hirayama, Y. & Yoshida, M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3618–3622 (1983).
- Cullen, B. R. *Meth. Enzym.* **152**, 692–704 (1987).
- Feinberg, M. R., Jarrett, R. F., Aldovini, A., Gallo, R. C. & Wong-Staal, F. *Cell* **46**, 807–817 (1986).

ACKNOWLEDGMENTS. We thank Dr Hiroaki Mitsuya for the 0.5- α monoclonal antibody, Dr Flossie Wong-Staal for the HTLV-I proviral pCR-1 plasmid, and Bonnie Kissell and Sharon Goodwin for manuscript preparation. Dr Madeleine Duc Dodon's work was supported in part by a grant from the Association pour la Recherche sur le Cancer, Villejuif, France.

FIG. 1 Topology of wild-type *E. coli* leader peptidase, and of the Lep' and Lep'-inv derivatives. The trypsin cleavage site in Lep'-inv is shown by the arrow. N terminal amino-acid sequences, and the nine-residue spacer in Lep' and Lep'-inv, are indicated (single-letter amino-acid code). Mutants of Lep were constructed by oligo-directed insertions and deletions using the Kunkel method¹⁴.



weight similar to the previously constructed $\Delta 4-50$ mutant⁷, consistent with a cleavage site in the P1 region. Radiosequencing of the protected fragment unambiguously shows that trypsin cleaves after Lys 37 (Fig. 3a) demonstrating that P1 is exposed to the periplasm and that P2 is in the cytoplasm in Lep'-inv.

To localize the N-terminus of Lep' and Lep'-inv, I took advantage of the fact that wild-type Lep has the N-terminal sequence f-Met-Ala-Asn-Met-Phe- (ref. 4). It is known from protein sequencing that the N terminus is blocked⁴; yet, all known cytoplasmic *E. coli* proteins starting f-Met-Ala- are unblocked and have lost the f-Met¹⁰. A possible explanation for this anomalous property of wild-type Lep is that the N terminus is facing the periplasm and hence protected from de-formylating and Met-removing enzymes. If this is indeed the case, one would expect mutant Lep-derivatives with a cytoplasmic N terminus to be unblocked and to lack the initial Met residue.

To test this idea, Lep' and Lep'-inv were labelled with [³⁵S]-Met and subject to radiosequencing. Like wild-type Lep, Lep' did not give a sequence, but Lep'-inv gave a strong signal in cycle 3, as would be expected if the f-Met had been removed (Fig. 3b). As a further test, the non-translocated $\Delta 5-22$ mutant⁸ lacking H1 also gave a strong signal indicating that f-Met had been removed (data not shown). Although indirect, these results suggest that wild-type Lep and Lep' have a periplasmic N

terminus, whereas the N terminus of Lep'-inv faces the cytoplasm.

Finally, as our previous statistical studies^{2,3} of polytopic membrane proteins led to the proposal that periplasmic regions shorter than 60-70 residues are translocated by a mechanism that is independent of the *secA* and *secY* genes, Lep' and Lep'-inv were transformed into the appropriate *secA*^{ts} and *secY*^{ts} strains. At the non-permissive temperature, translocation of Lep' and of the outer-membrane protein OmpA was blocked in both strains (Fig. 4). By contrast, the assembly of Lep'-inv was unaffected, confirming the prediction that translocation of the 25-residues-long P1 domain is independent of the SecA and SecY functions.

This study shows that the apolar regions H1 and H2 carry no topological information: the topology of Lep' is completely inverted when four lysines are placed at its N terminus. Although

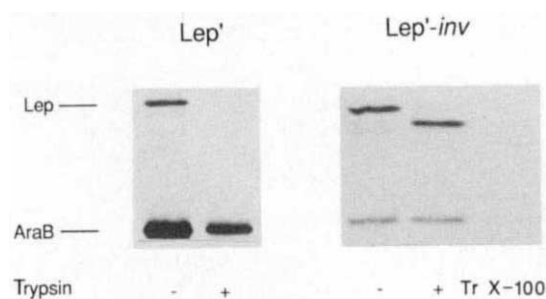


FIG. 2 Trypsin removes the P2 domain from Lep' and cleaves Lep'-inv in the P1 domain. The cytoplasmic protein, AraB, serves as a control for the integrity of the inner membrane. Mutants were expressed from the pING vector¹⁵ in strain HUM114 essentially as described previously⁸. Cells were grown in M9 minimal medium supplemented with ampicillin, 0.5% fructose and all amino acids except Met to early log phase, expression of the mutant protein was induced for 5 min by addition of arabinose to 0.2% final concentration, cells were labelled with [³⁵S]Met for 2 min, injected into ice-cold sucrose-EDTA buffer to permeabilize the outer membrane and incubated with trypsin (0.75 mg ml⁻¹) (or with trypsin + trypsin inhibitor + PMSF) for 1 h on ice. Where indicated, cells were lysed by vortexing in 2% Triton X-100 (Tr X-100) before trypsin treatment. Subsequently, soybean trypsin inhibitor (1 mg ml⁻¹) and PMSF (0.8 mg ml⁻¹) was added, samples were acid-precipitated, resuspended in 10 mM Tris/2% SDS, immunoprecipitated with antiserum to Lep and AraB and analysed by SDS-PAGE and autoradiography.

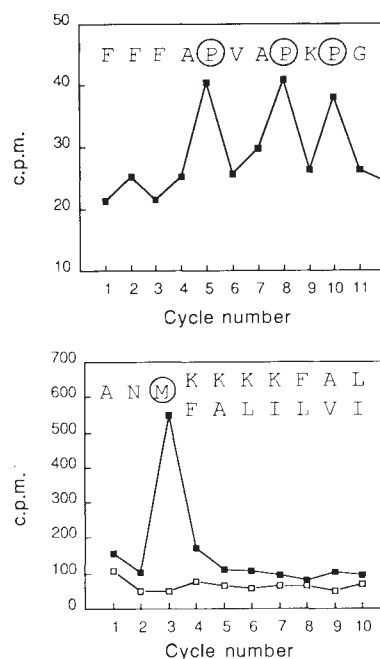


FIG. 3 N-terminal radiosequencing of the trypsin-resistant fragment from Lep'-inv (top) and of full-length Lep' and Lep'-inv (bottom, open and closed squares, respectively). The sequence following Lys 37 (top) and the N-terminal sequences following f-Met in Lep'-inv (bottom, upper) and Lep' (bottom, lower) are shown. Mutant proteins were expressed in strain MC1061 (0.5 ml) by induction with 0.2% arabinose for 5 min, labelled with 125 μ Ci [³H]Pro (top) or 60 μ Ci [³⁵S]Met (bottom), trypsin-treated (top only), acid-precipitated, resuspended in 10 mM Tris/2% SDS, immunoprecipitated with antiserum to Lep, and sequenced on an Applied Biosystems 470A protein sequencer.

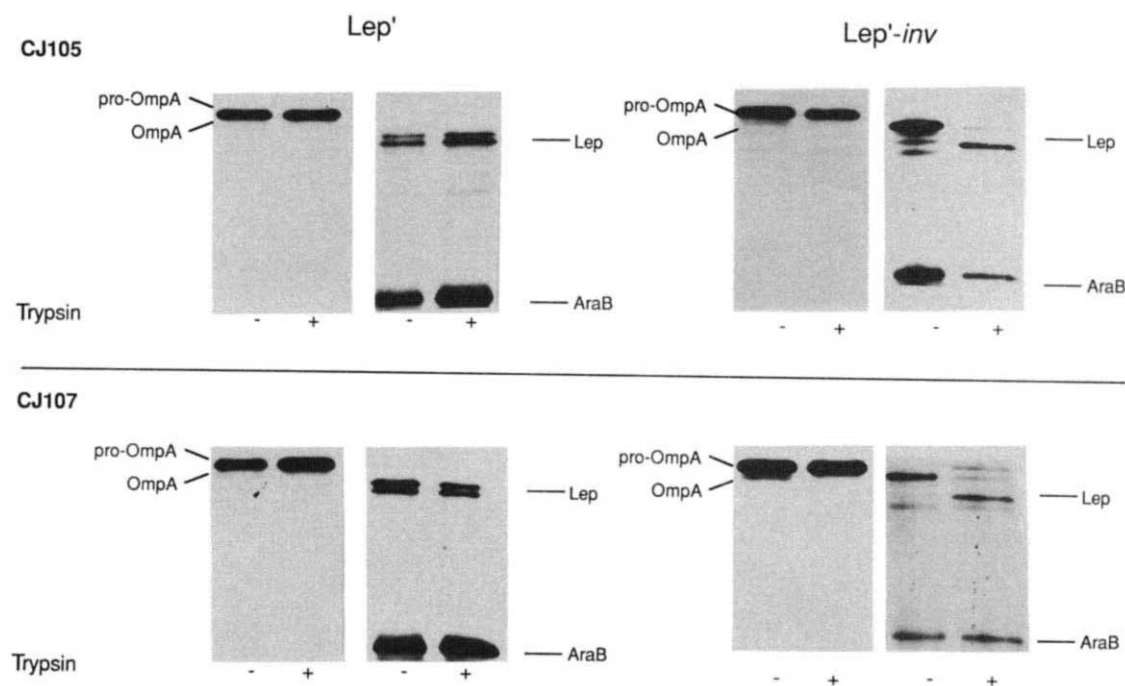


FIG. 4 Assembly of Lep' is SecA- and SecY-dependent whereas assembly of Lep'-inv is independent of the SecA and SecY proteins. pING plasmids encoding Lep' and Lep'-inv were transformed into strains CJ105 (secA^{ts}) and CJ107 (secY^{ts}) (ref. 9). Cells were grown at 30 °C to A₆₀₀ ≈ 0.1, and were then incubated at 42 °C for 3 h. Induction, labelling, and immunoprecipitation was done as in Fig. 1. Parallel immunoprecipitation of the outer-

membrane protein OmpA and its non-translocated precursor pro-OmpA was carried out as a control of the sec^{ts} phenotypes. The doublet bands in the Lep'-lanes probably result from limited intracellular proteolysis of the non-translocated protein; similar doublets have been observed in other Lep-mutants¹¹.

we have not yet tested the effects of negatively charged residues, statistical^{2,3} as well as experimental¹¹ studies suggest that their effect, if any, is weaker than that of positively charged residues. The results also strengthen the idea that short polar regions can be translocated by a sec-independent mechanism², irrespective of the size of the whole molecule¹². Longer regions apparently can only be translocated by the normal sec-dependent path-

way¹³. It seems, therefore, that efficient engineering of the transmembrane topology and mode of membrane assembly of polytopic membrane proteins can be based on the simple concepts of apolar membrane-spanning segments, polar flanking regions behaving according to the positive inside rule, and a critical length for sec-independent compared to sec-dependent assembly. □

Received 4 July; accepted 14 August 1989.

1. von Heijne, G. *Biochim. biophys. Acta* **947**, 307-333 (1988).
2. von Heijne, G. *EMBO J.* **5**, 3021-3027 (1986).
3. von Heijne, G. & Gavel, Y. *Eur. J. Biochem.* **174**, 671-678 (1988).
4. Wolfe, P. B., Wickner, W. & Goodman, J. *J. biol. Chem.* **258**, 12073-12080 (1983).
5. Moore, K. & Miura, S. *J. biol. Chem.* **262**, 8806-8813 (1987).
6. Dalbey, R. E. & Wickner, W. *Science* **235**, 783-787 (1987).
7. Dalbey, R. E., Kuhn, A. & Wickner, W. *J. biol. Chem.* **262**, 13241-13245 (1987).
8. von Heijne, G., Wickner, W. & Dalbey, R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 3363-3366 (1988).
9. Wolfe, P. B., Rice, M. & Wickner, W. *J. biol. Chem.* **260**, 1836-1841 (1985).

10. Flinta, C., Persson, B., Jörnvall, H. & von Heijne, G. *Eur. J. Biochem.* **154**, 193-196 (1986).
11. Laws, J. K. & Dalbey, R. E. *EMBO J.* **8**, 2095-2099 (1989).
12. Kuhn, A., Kreil, G. & Wickner, W. *EMBO J.* **6**, 501-505 (1987).
13. Kuhn, A. *Eur. J. Biochem.* **177**, 267-271 (1988).
14. Kunkel, T. A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 488-492 (1985).
15. Johnston, S., Lee, J.-H. & Ray, D. S. *Gene* **34**, 137-145 (1985).

ACKNOWLEDGEMENTS. Antisera were gifts from Dr Ulf Henning, Tübingen (OmpA), and from Dr Ross Dalbey, Ohio State University (Lep, AraB), who also provided technical guidance in the early phases of the project. This work was supported by the Swedish Natural Sciences Research Council and the Swedish National Board for Technical Development.

Unusual helical packing in crystals of DNA bearing a mutation hot spot

Youri Timsit, Eric Westhof, Robert P. P. Fuchs* & Dino Moras

Laboratoire de Cristallographie Biologique, and

*Groupe de Cancérogénèse et de Mutagénèse Moléculaire et Structurale, Institut de Biologie Moléculaire et Cellulaire, Centre National de la Recherche Scientifique, 15, rue René Descartes, 67084 Strasbourg Cedex, France

THE target sequence of the restriction enzyme *NarI* (GGCGCC) is a hot spot for the -2 frameshift mutagenesis (GGCGCC → GGCC) induced by the chemical carcinogens such as *N*-2-acetylaminofluorene^{1,2}. Of the guanine residues, all of which show equal reactivity towards the carcinogen, only binding to the 3'-most proximal guanine within the *NarI* site is able to trigger the frameshift event³. We selected the non-palindromic dodecamer d(ACCGGCGCCACA), whose sequence corresponds to the most mutagenic *NarI* site in pBR322 DNA, for X-ray structure analysis. Its molecular structure determined at 2.8 Å resolution reveals significant deviations from the structure of canonical B-form DNA, with partial opening of three G-C base pairs, high propeller twist values and sequence-dependent three-centred hydrogen bonds. This crystal structure shows a novel kind of packing in which helices are locked together by groove-backbone interactions. The partial opening of G-C base pairs is induced by interactions of phosphate anionic oxygen atoms with the amino group of cytosine bases. This provides a model for close approach of DNA molecules during biological processes, such as recombination.

The crystal structure presented here is that of the second dodecamer duplex crystallized in the B form, with a sequence and packing environment totally different from the d(CGCGAATTCGCG) dodecamer⁴. It provides a new example for the evaluation of sequence and packing effects on DNA structure. The sequence of the two complementary strands (Fig. 1) contains the six nucleotides of the *NarI* site in the centre of the duplex. The same sequence is present in the octamer d(GGGCGCCCC), which was crystallized in an A form⁵.

Crystallization conditions and preliminary crystallographic data are summarized in the legend to Fig. 1. Two points should be noted: (1) Ionic strength conditions used for crystal growth are quite similar to those generally used for growing B-form

DNA crystals, but the diffraction power of the crystals is strongly dependent on a correct stoichiometric magnesium/oligonucleotide ratio. The best crystals diffract to 2.8 Å when the magnesium/oligonucleotide ratio is 16. (2) There is a high sensitivity to temperature variation: lowering the temperature by a few degrees induces a phase transition with a slight rearrangement of the molecules and loss of symmetry, leading to a space group change from *R*3 to *P*31. Here we discuss the crystal structure in space group *R*3, presently refined to 15.3% for 1,101 significant reflections.

Mean values of helical parameters resemble those in canonical B-form DNA structures (Table 1). The minor groove width is 7 Å, in agreement with values found for mixed-sequence B-form DNA⁶. This value is close to that of the decamer containing two G-A mismatches (7.2 Å) (ref. 7) and to that of the phosphorothioate hexamer (6.4 Å) (ref. 8). But important deviations from standard B-form DNA⁹ are observed in the centre of the helix. Figure 2a shows the $3F_o - 2F_c$ map centred on the *NarI* site region of the duplex. It is clear that the C6-G7-C8 step has some very unusual conformational features. The direct interaction between an oxygen atom of a symmetry-related phosphate group and the amino group of cytosine C6 induces the partial opening of the C6-G19 base pair: the distance between N4(C6)-O6(G19) is 3.7 Å, 3.3 Å for N3(C6)-N1(G19), and 2.9 Å for O2(C6)-N2(G19). The C18-G7 base pair is disrupted in the same way, with a very high propeller twist (-30°) leading to unstacking of C18 from its two adjacent base planes. The high propeller twist value of G7 and G17 leads to close contact of their O6 atoms and to the partial opening of the C8-G17 base pair (with distances N4(C8)-O6(G17) equal to 3.7 Å, N3(C8)-N1(G17) to 3.2 Å, and O2(C8)-N2(G17) to 2.9 Å). This distorted base pair seems to be stabilized by a cross-strand, three-centred hydrogen bond (N4(C8)-O6(G16) distance = 3 Å) and by a partially hydrated magnesium ion bridging the O6 atoms. The overall conformation is further stabilized by three-centred hydrogen bonds. Figure 2b and c shows the high propeller twist value of each base pair, which allows the formation of non-coplanar hydrogen bonds. Distances between N4 (C8) and O6(G16), N4(C9) and O4(T15), N6(A10) and O6(G14) are all ~3 Å. These interactions are strongly sequence-dependent and concern half of the helix from C6 to A12. Three-centred hydrogen bonds have already been observed in other B-form DNA structures having A-tract sequences¹⁰⁻¹², A-T tract sequences⁶ and mismatched G-A pairs⁷. The results presented here extend the possibility of forming a network of three-centred hydrogen bonds to mixed sequences, including G-C base pairs, in a sequence-dependent manner.

These perturbations are probably induced and stabilized by

TABLE 1 Helical parameters of the dodecamer

	Roll (deg)	Tilt (deg)	Propeller twist (deg)	Twist (deg)	Rise (Å)
A1-G24	0	-4	-5	34	2.9
C2-G23	11	3	-20	35	3.7
C3-G22	3	-4	-26	38	3.5
G4-C21	-2	-4	0	33	3.1
G5-C20	1	4	-8	33	3.5
C6-G19	10	1	-7	35	3.7
G7-C18	2	-3	-30	36	3.2
C8-G17	3	1	-17	37	3.5
C9-G16	7	-2	-13	39	3.4
A10-T15	0	0	-14	28	3.3
C11-G14	6	1	-6	38	3.2
A12-T13	—	—	-17	—	—
Mean value	3.7	-0.6	-13.6	35.0	3.4

Helical parameters were calculated with the program NEWHELIX (distributed by R. E. Dickerson); nomenclature follows the Cambridge Conventions¹⁶.

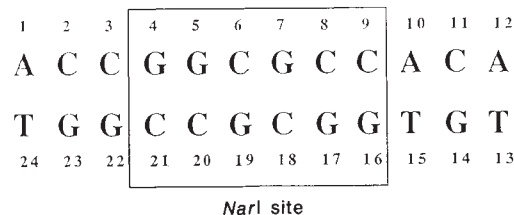


FIG. 1 Sequence of the dodecamer. The two complementary strands were synthesized by the phosphoramidites method. Single crystals were grown at 4 °C by the vapour diffusion technique from a solution containing 50 mM cacodylate at pH 6, 1 mM dodecamer duplex, 1.2 mM spermine tetrachloride and 18 mM magnesium acetate equilibrated against a reservoir of 45% 2,4-methylpentanediol. Crystals appeared in two or three days and grew during three weeks to reach 500 × 200 × 200 μm. The best diffracting crystals (2.8 Å) are produced at a magnesium/oligonucleotide stoichiometric ratio >16:1. For lower ratios, crystals only diffract at low resolution. Space group is *R*3, with cell dimensions: *a* = *b* = 65.9 Å, *c* = 47.1 Å, and one duplex per asymmetric unit.

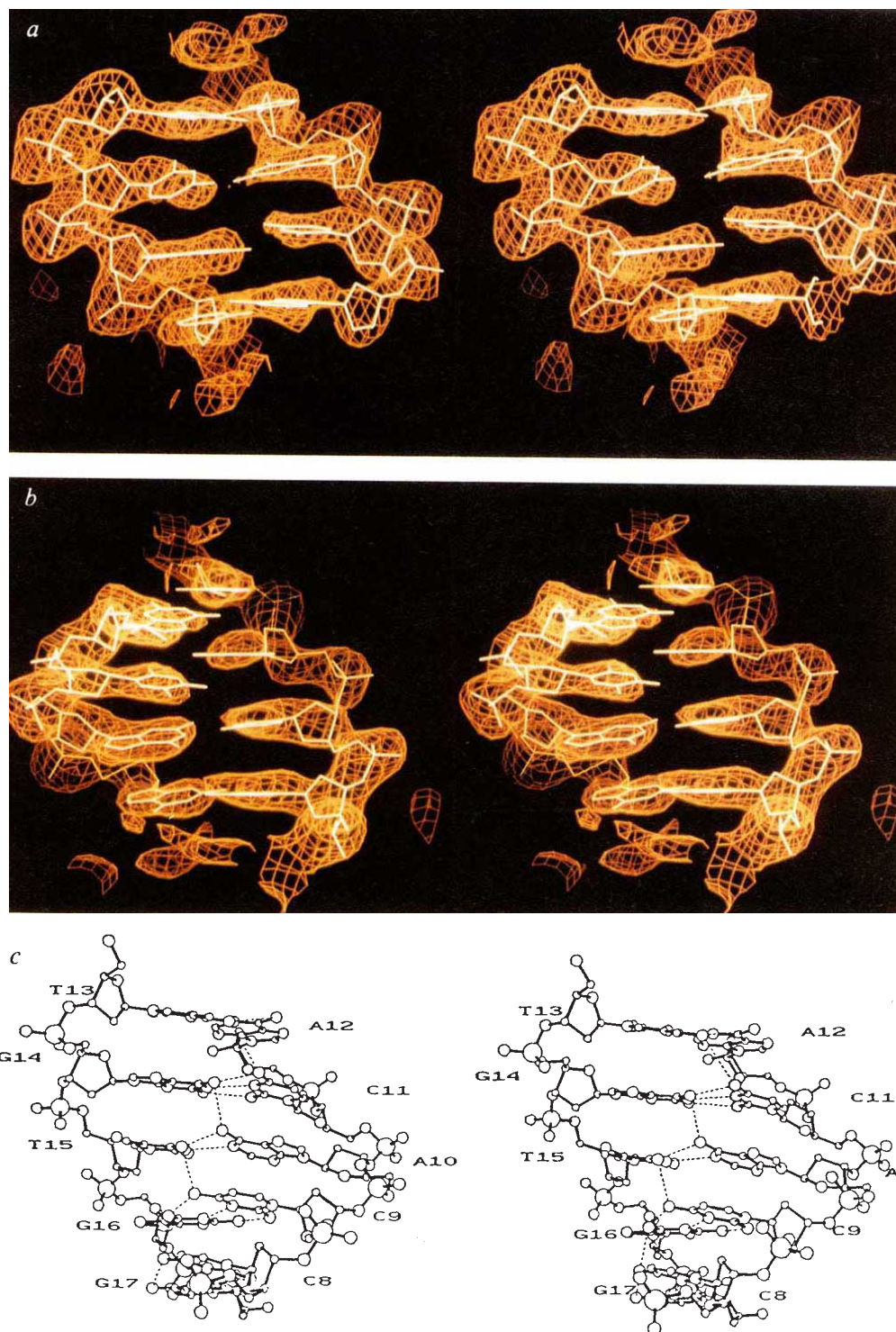
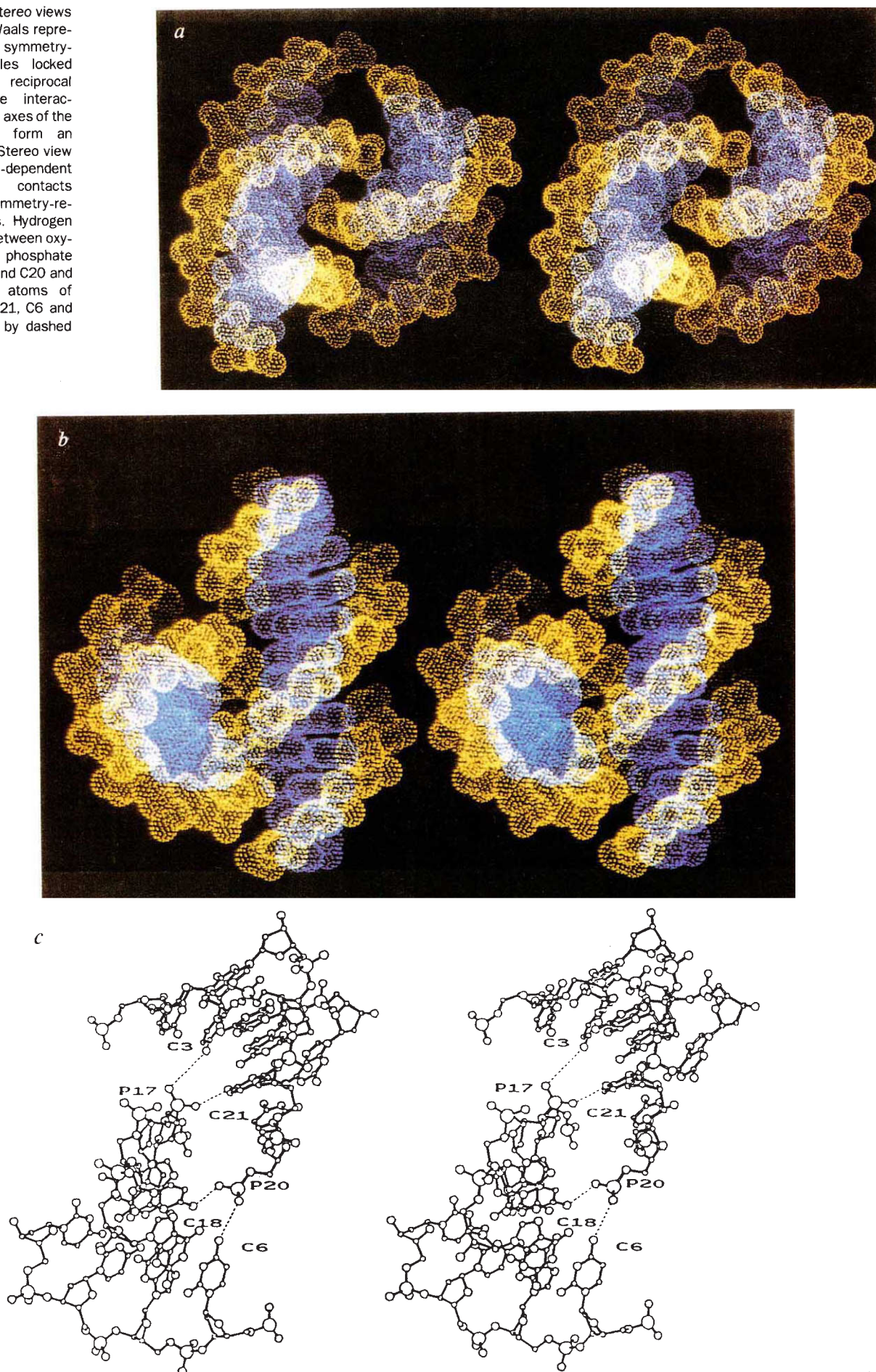


FIG. 2 *a*, $3F_0-2F_c$ map centred on the *NarI* site region of the duplex, from the G5-C20 pair (bottom) to the C8-G17 pair (top). This stereo view of the major groove shows the unstacked C18 base with high propeller twist (30°), which leads to the disruption of the G7-C18 pair. *b*, $3F_0-2F_c$ map showing high propeller twist and three-centred hydrogen bonds, from the C8-G17 pair (bottom) to the C11-G14 pair (top). *c*, Stereo drawing of the sequence-dependent array of three-centred H-bonds. Cross-strand distances are N4(C8) to O6(G16), 3 Å; N4(C9) to O4(T15), 2.6 Å; N6(A10) to O6(G14), 2.7 Å, N4(C11) to O4(T13), 3.3 Å.

METHODS. Data were collected at 0°C on a Nicolet detector, using three crystals. 1,584 independent reflections with R_{sym} on intensity equal to 7.7% were obtained above the 1σ level. The structure was solved by molecular replacement using the program ULTIMA (ref. 17), with a standard B-form DNA adapted from the fibre model of Arnott⁹, with the proper sequence.

Rigid body refinement led to an R factor of 21% for the 25–10 Å data. The refinement was continued with the molecular dynamic program GROMOS (ref. 18) using 10–3.5 Å data above the 4σ level (I_{obs}), to reach R factor of 13%. Resolution was gradually extended to 2.8 Å for further refinement with the NUCLSQ program¹⁹. No restraints were placed on base pairing. From the two possible orientations around the pseudo 2-fold axis of the helix, we kept the one that gave an unambiguously better fit on all three terminal base pairs. Further stereochemical arguments have confirmed this choice. Final R factor with 1,101 reflections above the 3σ level (I_{obs}) reached 15.3% and 19.6% with the 1,584 reflections above 1σ (I_{obs}). The cluster of 30 strong reflections in 3.4 Å area was eliminated from the data above the 3σ level used for $3F_0-2F_c$ map calculation. Twenty water molecules and one hydrated magnesium could be located on difference Fourier maps. Final coordinates will be sent to the Brookhaven Protein Data Bank.

FIG. 3 *a* and *b*, Stereo views of the van der Waals representation of two symmetry-related molecules locked together by reciprocal groove-backbone interactions. The helical axes of the two molecules form an angle of 60° . *c*, Stereo view of the sequence-dependent intermolecular contacts between two symmetry-related molecules. Hydrogen bonds formed between oxygen atoms of phosphate groups of C17 and C20 and the amino N4 atoms of cytosines C3, C21, C6 and C18 are shown by dashed lines.



packing forces, and so crystallization selected one structural state among the various conformations possible for this sequence. Our crystal structure could be viewed as a snapshot picture of a CGC motif in a 'transition state' before base-pair opening. Two types of intermolecular contacts are present in the crystal packing: the end-to-end stacking of duplex molecules forming a pseudo-continuous helix, and the groove-backbone interactions, by which symmetry-related molecules are locked together as a result of shape and electrostatic complementarity between the backbone and the helix grooves (Fig. 3a and b). Direct interactions involving the N4 atoms of cytosines of one molecule and oxygen atoms from phosphate groups of a second molecule form the shortest intermolecular contacts. Figure 3c shows a detailed view of the sequence-specific packing contacts. It can be seen that alternate cytosines point in the major groove and interact with the phosphate group in a sequence-dependent manner: in the major groove of one molecule, N4 atoms of C18 and C6 interact with O1P and O2P atoms of C20 of a symmetry-related molecule. In the major groove of the second molecule, the same interaction is reversible between the N4 atoms of C3 and C21 and the oxygen atoms of the phosphate group of G17. The distances between the amino group and the anionic oxygen atoms of the phosphate groups are all ~ 3 Å. Indirect contacts involving both a magnesium ion and a water molecule bridge the O6 of G17 and the O1P atom of G19.

The new type of oligonucleotide packing observed in this structure, where helices are locked together in a sequence-dependent manner, presents an interesting analogy with the major groove- α -helix interactions found in repressor-operator complexes^{13,14}, in which important deviations of base-pair geometry with three-centred hydrogen bonds are also found. Furthermore, it provides a new model for close approach between helices as found, for example, during recombination, and extends the concept of specificity of molecular recognition to DNA-DNA interactions. Also, there might be a similar magnesium requirement influencing the stability of this crystal packing and the intermediate stages of recombination¹⁵. Alternate NH_2 groups in the major groove and magnesium ions might play an important part in specifying and stabilizing the Holliday junction. This crystal structure also illustrates how the introduction of a phosphate group in the major groove could provide an easy way to initiate denaturation through stabilization of open base pairs.

In addition, it is tempting to conclude that the pronounced deviations from the standard B-form DNA structure in the centre of the *NarI* site are related to the mutation-hot-spot properties of this sequence. Indeed, these peculiar structural properties may favour the induction of a specific conformational change when N-2-acetylaminofluorene is bound to G7, the only binding event triggering mutation³. □

The α -lytic protease pro-region does not require a physical linkage to activate the protease domain *in vivo*

Joy L. Silen & David A. Agard

Department of Biochemistry and Biophysics, and
The Howard Hughes Medical Institute, University of California,
San Francisco, California 94143, USA

α -LYTIC protease, an extracellular serine protease of *Lysobacter enzymogenes* 495, is synthesized as a pre-pro-protein¹. Previously it has been shown that when expressed in *Escherichia coli*, the protein is autocatalytically processed in the periplasmic space, and that the functional protease domain accumulates extracellularly². Engineered proteins lacking the 166 amino-acid pro-region were enzymatically inactive and remained cell-associated². By independently expressing the pro- and protease domains *in vivo*, evidence is provided here that direct covalent linkage is not required for production of active protease. We postulate that the pro-region acts as a template to promote the folding of the protease domain into an active configuration. Our results, combined with recent experiments on the evolutionarily unrelated subtilisin E (ref. 3), suggest that the ability of the pro-region of these bacterial proteases to facilitate folding of their protease domains is not a curiosity of a single system, but may reflect a general property of extracellular bacterial serine proteases.

On the basis of denaturation and renaturation experiments on a wide variety of small proteins⁴⁻⁶, it seems that small, single-subunit, globular proteins fold spontaneously to reach a minimum energy state. For example, Anfinsen demonstrated that ribonuclease A could be unfolded and then refolded with full restoration of activity⁴. Unexpectedly, two small bacterial serine proteases, subtilisin (J. Wells, personal communication) and α -lytic protease (J. Richards, personal communication) could not be refolded. Experiments in both systems^{2,7} indicated that the pro-region, which is covalently attached to, but only transiently associated with the protease domain, is necessary for correct folding of these enzymes. Several different models can be proposed for the role of the pro-region in folding the protease domain (Fig. 1). In models 1 and 2, the properly folded protease is considered to be at an energy minimum, but is kinetically inaccessible. That is, there are one or more non-productive folding pathways leading to improperly folded and possibly aggregated protein that have a lower transition-state energy and higher ground-state energy than the productive pathway. The pro-region would then facilitate production of active protease, either by stabilizing the transition state for the correct pathway (Fig. 1, model 1) or by destabilizing the non-productive pathways (Fig. 1, model 2). Unfortunately, it is not yet possible to distinguish between these two models as this requires a detailed kinetic analysis using purified components. Model 3 (Fig. 1a, b), however, which depends on the energy liberated from cleavage of the precursor, can be tested. In this model, the functional protease is not at an energy minimum (the misfolded forms on the non-productive pathway have lower energies), but is kinetically trapped. In such a model, the pro-region would act as in model 1 to lower the kinetic barrier, but in addition, the energy released by cleaving the precursor is required to shift the overall equilibrium in favour of the productive pathway. Once formed, the active high-energy protease would be unstable in the presence of functional pro-domain (Fig. 1b). To avoid unfolding active protease and subsequent accumulation in the lower-energy misfolded state, the pro-region must be physically separated from the protease or inactivated.

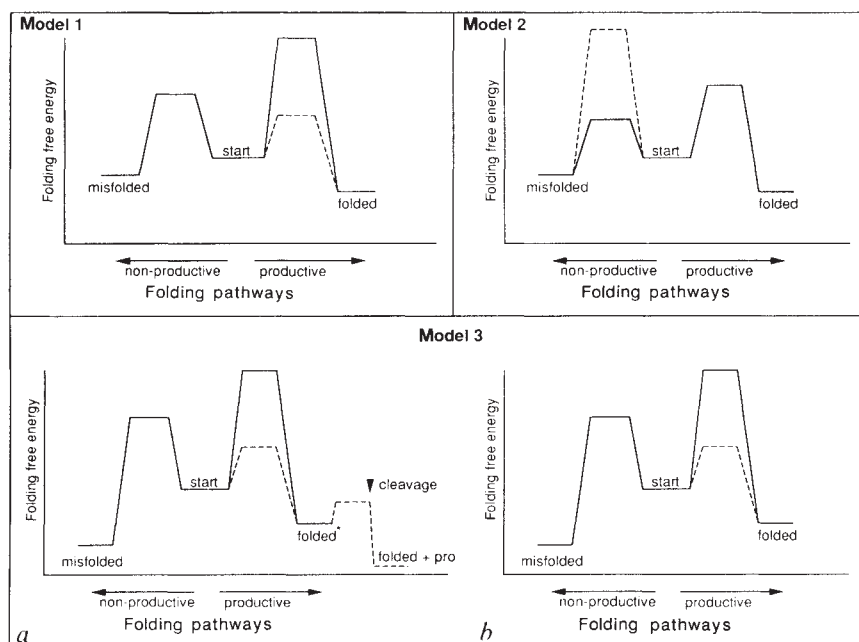
To test if the physical linkage and its subsequent cleavage

Received 30 June; accepted 21 August 1989.

1. Fuchs, R. P. P., Schwartz, N. & Daune, M. *Nature* **294**, 657-659 (1981).
2. Koffel-Schwartz, N. et al. *J. molec. Biol.* **177**, 33-51 (1984).
3. Burnouf, D., Koehl, P. & Fuchs, R. P. P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4147-4151 (1989).
4. Wing, R. et al. *Nature* **287**, 755-758 (1981).
5. Rabinovich, D., Haran, T., Eisenstein, M. & Shakked, Z. *J. molec. Biol.* **200**, 151-161 (1988).
6. Yoon, C., Privé, G., Goodsell, D. & Dickerson, R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6332-6336 (1988).
7. Privé, G. et al. *Science* **238**, 498-504 (1987).
8. Cruse, W. et al. *J. molec. Biol.* **192**, 891-905 (1986).
9. Arnott, S. & Hukins, D. *Biochem. biophys. Res. Commun.* **47**, 1504-1509 (1972).
10. Nelson, C., Finch, J., Luisi, B. & Klug, A. *Nature* **330**, 221-226 (1987).
11. Coll, M., Frederick, C., Wang, A. & Rich, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8385-8389 (1987).
12. DiGabriele, A., Sanderson, M. & Steitz, T. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1816-1820 (1989).
13. Aggarwal, A., Rodgers, D., Drott, M., Ptashne, M. & Harrison, S. *Science* **249**, 899-907 (1988).
14. Jordan, S. R. & Pabo, C. O. *Science* **242**, 893-899 (1988).
15. Duckett, D. et al. *Cell* **55**, 79-89 (1988).
16. Embo Workshop *EMBO J.* **8**, 1-4 (1989).
17. Rabinovich, D. & Shakked, Z. *Acta Crystallogr.* **A40**, 195-200 (1984).
18. Fujinaga, M., Gros, P. & Van Gunsteren, W. F. *J. appl. Crystallogr.* **22**, 1-8 (1989).
19. Westhof, E., Dumas, P. & Moras, D. *J. molec. Biol.* **184**, 119-145 (1985).

ACKNOWLEDGEMENTS. We are grateful to A. Hoeft for the synthesis of the dodecamer. We thank A. Bilwes, B. Chevrier, J. Colin, A.-C. Dock-Bregeon, B. Rees and J.-C. Thierry for discussion. Y.T. is a recipient of an IRSIA fellowship.

FIG. 1 Three models for the role of the pro-region in folding the protease domain. In all cases the folding process is considered to be a competition between a productive folding pathway that leads to the correctly folded protease, and a non-productive pathway that leads to misfolded and/or aggregated inactive protein. The free energy of folding from the point of view of the protease domain is shown in the absence (solid lines) and presence (dashed lines) of the pro-region. In model 1, the pro-region is seen to stabilize the transition state for folding, whereas in model 2 the pro-region destabilizes the transition state for the competing non-productive pathway. Models 1 and 2 are independent of the covalent attachment of the pro-region and thus, for simplicity, the energetics of the pro-region cleavage are not included. Model 3, depicting a kinetically trapped, higher energy folded state, is shown with either a covalently attached pro-region (panel a) or with a non-linked pro-region (panel b). The energy of bond cleavage can be used to shift the overall folding equilibrium in favour of the folded configuration only if the pro-region is covalently attached (a versus b). In model 3 there must be a folded (or mostly folded) intermediate that is still covalently linked to the pro-domain (folded* in panel a), which after cleavage results in the final, active form of the protease.



plays a direct part in the folding process (model 3), it was necessary to independently express the pro- and protease domains within the periplasmic space. Expression systems with signal sequences were employed because the protease has three disulphide bonds and thus requires an oxidizing environment to fold properly. The expression construct, pCOMP5 (Fig. 2), uses the *phoA* promoter and signal sequence fused to the protease domain² and the *lpp/lac* promoter and *ompA* signal sequence⁸ fused to the pro-region. This construct allows induction of the pro-region by isopropyl- β -D-thiogalactoside (IPTG) and the protease domain by phosphate starvation. Only the dual conditions of phosphate starvation and IPTG induction would allow complementation and, therefore, accumulation of the active enzyme.

The results of the induction experiment are shown in Fig. 3. As expected, independent induction of either the protease or pro-domains results in little or no activity. The low level of

proteolytic activity seen in the low-phosphate control culture, in which only the protease domain is induced, probably results from the basal expression of the pro-region by the *lpp/lac* promoter⁸. In the cultures containing low phosphate plus IPTG, in which both promoters are induced, α -lytic protease activity reaches levels more than 30 times that of the low-phosphate control culture. As the accumulation of active α -lytic protease occurs after the addition of IPTG and follows the induction of cellular alkaline phosphatase, it is clear that co-induction is required for α -lytic protease production.

These experiments, and the *in vitro* experiments with subtilisin⁴, demonstrate the independence of the folding process from the direct linkage between the pro- and protease domains, thereby eliminating the kinetic trapping model (Fig. 1, model 3). By uncoupling the irreversibility of bond cleavage from the folding process, the thermodynamics of folding of the protease domain can be more accurately considered, and more easily

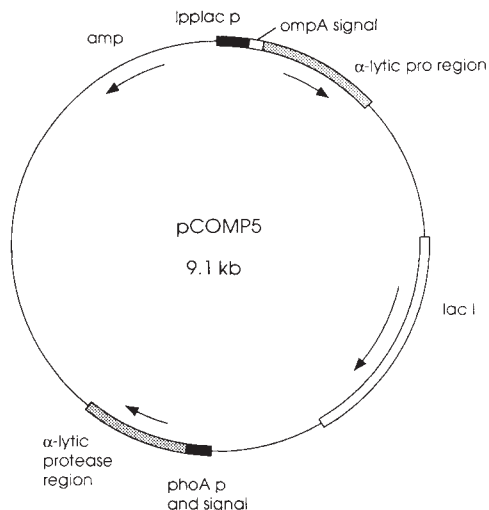


FIG. 2 The pCOMP5 plasmid (9.1 kilobases). This pIN-III-ompA2 derivative⁸ places the pro-region of α -lytic protease under the control of the IPTG-inducible *lpp/lac* promoter and downstream of the *ompA* signal sequence. The protease domain of the α -lytic protease gene is fused to the *phoA* promoter and signal sequence; its expression depends on depletion of phosphate in the medium.

METHODS. The pro-region was originally isolated as a filled-in *EaeI* fragment from pDA13 (ref. 1) and cloned into a modified pDBR2 cut with *SmaI* and *EcoRI* (blunted). This construct (pPRO1) established *phoA*-promoter controlled expression and placed an in-frame stop codon six amino acids from the end of the pro-region. A filled-in *EaeI*-*BalI* fragment from pPRO1 was cloned into the filled-in *EcoRI* site of pINIII-ompA2, which placed it in frame with the *ompA* signal sequence (pPRO3). The coding region for the protease domain of α -lytic protease was added to pPRO3 at a unique blunted *StyI* site by transferring in a filled-in *NotI*-*NarI* fragment from pMALP4 (a derivative pMALP2 (ref. 2). Orientation of the protease-encoding fragment was established by restriction enzyme analysis. Constructs were transformed into *E. coli* strain HB101 (*recA*⁻) for expression.

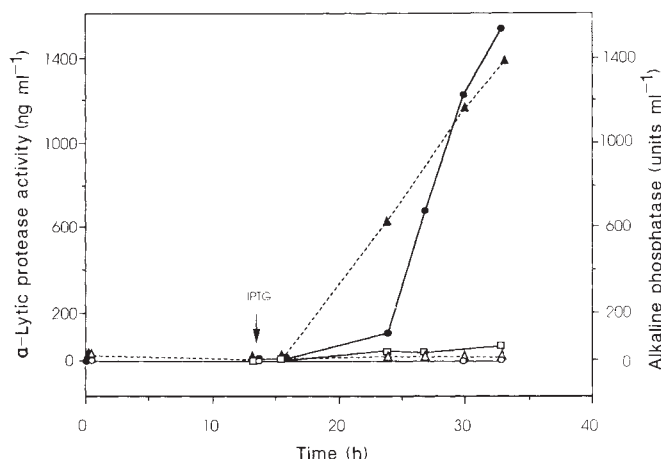


FIG. 3 Time course of α -Lytic protease and alkaline phosphatase production in cultures of HB101/pCOMP5. As the exact time of phosphate depletion is uncertain, levels of chromosomal alkaline phosphatase (--- $\blacktriangle$ ---) were monitored to follow the time course of the induction of the *phoA* promoter. Alkaline phosphatase levels under non-inducing conditions (high phosphate plus IPTG) remained at baseline levels (--- $\triangle$ ---). α -Lytic protease production was also monitored as a function of time: significant levels of activity are seen only when both the pro- and protease regions are co-expressed (compare low phosphate plus IPTG (--- $\bullet$ ---) with phosphate plus IPTG (--- $\circ$ ---) or low phosphate alone (--- $\square$ ---)).

METHODS. Cells grown overnight in high-phosphate medium were concentrated, resuspended in low-phosphate medium, and diluted 1:100 in 250 ml low phosphate (0.1 mM K_2HPO_4/KH_2PO_4 , pH 7.2) MOPS medium, or directly diluted 1:100 in 250 ml of high phosphate (20 mM K_2HPO_4/KH_2PO_4 , pH 7.2) MOPS medium to an identical optical density at 600 nm ($OD_{600} = 0.04$). Four 50 ml aliquots of the low-phosphate culture were taken into 500 ml shaker flasks for duplicates of the low-phosphate and low-phosphate-plus-IPTG cultures; two 50 ml aliquots of the high-phosphate culture were used for the high phosphate duplicates. Cultures were grown with shaking (200 r.p.m.) at 22 °C. After 14 h ($OD_{600} = 1.0$), IPTG (2 mM final concentration) was added to the duplicate high-phosphate cultures and to one pair of the low-phosphate cultures. Levels of alkaline phosphatase (both from the cells and the medium), and α -lytic protease (medium only) were determined as previously described², using *p*-nitrophenyl phosphate and suc-Ala-Pro-Ala-*p*-nitroanilide, respectively, as substrates (α -lytic protease purified from *Lysobacter enzymogenes* has 188 units mg^{-1} ; 1 unit hydrolyses 1 μ mol substrate per min). Data shown is the average of the duplicate cultures, except for the low-phosphate culture alkaline phosphatase induction, which is the average of all four low-phosphate cultures, plus and minus IPTG.

investigated experimentally. Because active α -lytic protease displays remarkable temperature stability⁹, resistance to autolysis and a high degree of structural rigidity¹⁰, there is probably a high-energy barrier close to the folded state that provides the active enzyme with its high stability, as in model 1. Apparently, this barrier is too high to be traversed unaided. Thus, the pro-region can be thought of as a template that preferentially binds to, and stabilizes, the transition state for folding, just as an enzyme binds and stabilizes the transition state for catalysis of a reaction.

It is quite remarkable that these two classes of serine protease⁵, which are unrelated in structure and sequence, have, through convergent evolution, developed both a spatially identical Asp-His-Ser catalytic triad and what appears to be a similar folding mechanism. Molecular analysis of the genes for a wide variety of secreted proteases from both Gram-positive and Gram-negative bacteria¹¹⁻¹⁸ reveals that these enzymes are also synthesized in precursor form. The widespread use of pro-regions by extracellular bacterial serine proteases indicates that these domains must have an extremely important function.

The folding activity of these pro-regions is reminiscent of the recently discovered class of proteins, called chaperonins, which facilitate the folding and assembly of multi-subunit proteins in chloroplasts¹⁹, mitochondria²⁰ and of bacteriophage in *E. coli*¹⁹. Although the chaperonins are capable of aiding the folding of a number of proteins, these extracellular bacterial serine proteases seem to have evolved their own specialized folding aids which are covalently attached. Further examination of the structure and mechanism of the α -lytic protease pro-region may provide new insights into the function of other more generalized folding and assembly systems. $\square$

Received 10 May; accepted 14 August 1989.

1. Silen, J. L., McGrath, C. N., Smith, K. R. & Agard, D. A. *Gene* **69**, 237-244 (1988).
2. Silen, J. L., Frank, D., Fujishige, A., Bone, R. & Agard, D. A. *J. Bact.* **171**, 1320-1325 (1989).
3. Zhu, X., Ohta, Y., Jordan, F. & Inouye, M. *Nature* **339**, 483-484 (1989).
4. Anfinsen, C. B. *Proc. natn. Acad. Sci. U.S.A.* **47**, 1308-1314 (1961).
5. Creighton, T. E. *J. phys. Chem.* **89**, 2452-2459 (1985).
6. Ikai, A., Fish, W. & Tanford, C. J. *J. molec. Biol.* **145**, 265-280 (1973).
7. Ikemura, H., Takagi, H. & Inouye, M. *J. biol. Chem.* **262**, 7859-7864 (1987).
8. Ghayeb, J. et al. *EMBO J.* **10**, 2437-2442 (1984).
9. Kaplan, H., Symonds, V. B., Dugas, H. & Whitaker, D. R. *Can. J. Biochem.* **48**, 649-658 (1970).
10. Fujinaga, M., Delbaere, L. T. J., Brayer, G. D. & James, M. N. G. *J. molec. Biol.* **183**, 479-502 (1985).
11. Jacobs, M., Eliasson, M., Uhlen, M. & Flock, J. I. *Nucleic Acids Res.* **13**, 8914-8926 (1985).
12. Shimada, H. et al. *Biotechnology* **2**, 75-85 (1985).
13. Stahl, M. L. & Ferrari, E. *J. Bact.* **158**, 411-418 (1984).
14. Takagi, M., Imanaka, T. & Aiba, S. *J. Bact.* **163**, 824-831 (1985).
15. Vasantha, N. et al. *J. Bact.* **159**, 811-819 (1984).
16. Wells, J. A., Ferrari, E., Henner, D. J., Estell, D. A. & Chen, E. Y. *Nucleic Acids Res.* **11**, 7911-7925 (1983).
17. Yang, M. Y., Ferrari, E. & Henner, D. J. *J. Bact.* **160**, 15-21 (1984).
18. Henderson, G., Krygsman, P., Liu, C. J., Davey, C. C. & Malek, L. T. *J. Bact.* **169**, 3778-3784 (1987).
19. Hemmingsen, S. et al. *Nature* **333**, 330-334 (1988).
20. McMullin, T. W. & Hallberg, R. L. *Molec. cell. Biol.* **8**, 371-380 (1988).

ACKNOWLEDGMENTS. We thank Roger Bone, Amy Fujishige and Hiten Madhani for comments and M. Inouye for providing the pIN-III-ompA2 expression plasmid. This work was supported by the NSF Presidential Young Investigator Program and the Howard Hughes Medical Institute.

Copies of articles from this publication are now available from the UMI Article Clearinghouse.

For more information about the Clearinghouse, please fill out and mail back the coupon below.

UMI Article Clearinghouse

Yes! I would like to know more about UMI Article Clearinghouse.

I am interested in electronic ordering through the following system(s):

- ☐ DIALOG/Dialorder ☐ ITT Dialcom
☐ OnTyme ☐ OCLC ILL Subsystem

☐ Other (please specify) _____

☐ I am interested in sending my order by mail.

☐ Please send me your current catalog and user instructions for the system(s) I checked above.

Name _____

Title _____

Institution/Company _____

Department _____

Address _____

City _____ State _____ Zip _____

Phone (_____) _____

Mail to: University Microfilms International
 300 North Zeeb Road, Box 91 Ann Arbor, MI 48106

Modelling in peptide design

B. A. Jameson

Powerful molecular modelling software can be used to visualize the tertiary structures of biologically significant proteins, with important implications for rational drug design.

THE increasing database of crystallized protein structures combined with the increasing user-friendliness of powerful computer algorithms designed to calculate folding patterns of proteins has created an environment in which the empiric art of peptide and drug design is being replaced with designs based on shape. There is a current trend towards protein engineering where, for example, the goal is to change a protein's shape in order to alter its biological activity. Altering activity, however, requires extensive assumptions about a protein's mechanistic details that often extend far beyond available information. The goal of isosteric (same-shape) analogue design, on the other hand, is to mimic, not alter, the shape of a protein. This requires less guess-work and is, consequently, a more feasible approach at present.

In order to design an isosteric analogue of a protein, a template structure must, obviously, be available. Although the number of protein structures being deposited in the protein databank at Brookhaven National Laboratory is increasing at an astounding rate, the total number of available structures is still limited. Molecular modelling based on homology is a powerful means of extending this structural database. Recently, a high-resolution crystallographic structure of HIV-1 protease became available¹. Before the availability of this structure, there were two published reports of molecular models of this protein based on the assumption that backbone folding patterns within a given family of proteases is conserved. In one case a distantly related pepsin structure was used as a folding template², and in the other case a more closely related retroviral protease was used³. In both cases, the high-resolution crystal structure showed that these predicted structures were remarkably accurate. In this manner the structural database can be expanded to include evolutionarily related protein family members and, thus, increase the general applicability of rationally designed synthetic peptides.

Modelling parameters

There are currently five major software packages available for molecular modelling (BioDesign, Inc., Pasadena, California; Biosym Technologies, San Diego, California; Tripos, St. Louis, Missouri; Polygen, Waltham, Massachusetts; and Chemical Design Ltd, Oxford, UK). All of the molecular modelling packages operate

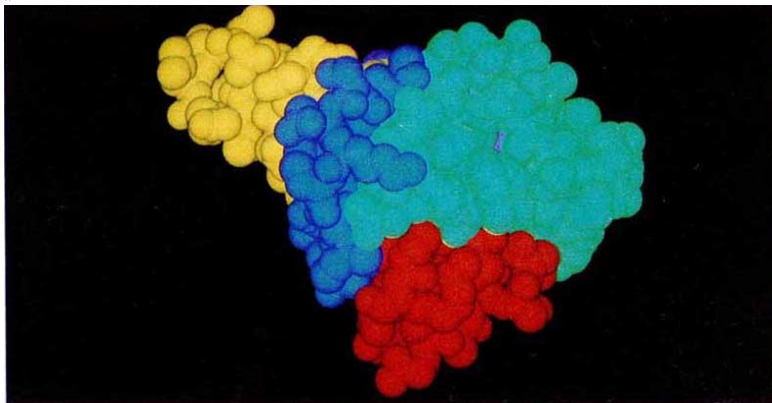


FIG. 1 Molecular model of the V₁ domain of the CD4 protein (residues 1–104). The model was constructed using an immunoglobulin light chain variable domain homodimer (Brookhaven database code: REI, Epp¹³) as the backbone template and using the molecular modelling package Biograf (BioDesign, Inc., Pasadena, California).

by using three basic parameters: a descriptive energy field, an algorithm for performing molecular mechanics and an algorithm for performing molecular dynamics^{4,5}.

The first component of a modelling package, the energy field, involves a set of parameter tables that are used to create mathematical descriptions of atoms and their potential interactions with one another based on averaged values determined from known structures. One of these parameter tables is an atom data file. This file contains descriptions of individual atoms, such as their partial charge, mass and type. Such tables provide the basic information necessary appropriately to represent the stereochemical properties of each atom. For example, carbon-carbon interactions with *sp*³ and *sp*² hybridizations need to be distinguished (single bond versus double bond) in terms of geometry, bond strength, ideal strength and other parameters. A bond description file is used in depicting the covalent structure of the molecule on the basis of the number and type of covalent bonds occurring between atom pairs. This type of table also includes descriptions of the 'spring-like' properties of various bonds and favourable and unfavourable torsional rotations. Finally, there is a table used in the consideration of nonbonded interactions such as van der Waals and electrostatic attractions.

The second and third components of the modelling package involve energy minimization functions, molecular mechanics, and energy-dependent molecular motion simulations and molecular dynamics. The theoretical goal of molecular mechanics is to calculate a local energy minimum for a given structure. Due to the number of

independent variables and quantum mechanical considerations, such calculations are not truly feasible and must be simplified. The algorithms used in treating potential energy functions of a structure through molecular mechanics serve to reduce or eliminate steric conflicts and adjust bond lengths and bond angles to approximate their ideal values. Generally, this technique will not produce structures grossly different from the initial structure. Molecular dynamic calculations are used to alter backbone structures. These programs rely on many of the same basic parameters as mechanics, except that temperature, pressure and atomic velocities are included in the calculations as a function of time. Typically, these calculations are performed repetitively in 0.002 ps stepwise intervals.

There are quite a few practical restrictions placed upon the modelling algorithms. One of the greatest limitations is that most of the spontaneous protein folding events in nature occur on a nanosecond time scale. Outside of calculations performed on Cray supercomputers (which are also time-restricted), most computational times are limited to under 100 ps. This means that a substantial body of information should be known about the protein's folding properties, such as its backbone folding architecture, before attempting to create a protein model. There exists a commonly held belief that the computer modelling algorithms are capable of taking a linear amino acid sequence and predicting a meaningful tertiary structure. Such is not the case. Understanding the basic limitations of these folding techniques is critical to producing a meaningful structure.

Our laboratory has been studying the

binding events that occur between the envelope protein (gp120) of the virus causing AIDS (HIV) and its cellular receptor, the CD4 protein. It has recently been shown that the first 104 amino acids of the CD4 protein are sufficient to account for the high-affinity interaction observed in binding to the gp120 protein⁶. This region of the CD4 protein is a part of the immunoglobulin superfamily of proteins^{7,8}. Although a crystal structure for this protein is not yet available, its relationship to the immunoglobulin family of proteins and the fact that it is a small, well-defined region make it an ideal candidate for molecular modelling. Our computer-generated model of this protein is shown in Fig. 1. The regions that correspond to antigen-binding domains on an immunoglobulin protein are highlighted (blue — CDR1, red — CDR2, yellow — CDR3). The CDR2-like region projects well away from the surface of the protein relative to the parent immunoglobulin structure. Site-directed mutagenesis experiments suggest that this region is directly involved in binding to the gp120 envelope protein⁹⁻¹¹. Our model agrees well with this finding.

We have now used the modelled CD4 protein as a template in the design of conformationally restricted synthetic peptides. This has been done by introducing artificially positioned disulfide bridges to force conformational folding of the synthetic peptide. While previously designed peptides that encompassed the linear sequence of the CDR2-like domain failed to show any effect, our crude isosteric analogues appear to have biological activity¹².

The work described above serves as an illustrative example of the limitation and potential power of the use of molecular modelling techniques. The protein, whose structure is unknown, belongs to a larger evolutionary family of proteins that include members with established structures. Biological data has defined a discrete subdomain of the protein for analysis, thus limiting the scope of the modelling problem. Perhaps most importantly, there exist biological assays to test the validity of the structural model. We feel that this type of approach can be successfully applied to a variety of different systems, and it may be possible to create powerful new peptide analogues for understanding structure/function relationships as well as reagents that may have useful applications as therapeutic drugs. □

Bradford A. Jameson is at the Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine, 3420 North Broad Street, Philadelphia, Pennsylvania 19140, USA. For more information, fill in reader service number 100.

1. Wlodawer, A. *et al. Science* **245**, 616 (1989).
2. Pearl, L. & Taylor, W. *Nature* **329**, 351 (1987).

Creative computing

A robotic titration workstation and 3-D molecular modelling software are a selection of what next week's Scientific Computing and Automation Conference in Philadelphia, Pennsylvania, has to offer.

SUN Microsystems, Inc. are emphasizing the compactness, high-performance capabilities and affordability of the SPARCstation 1 **desktop computing system**, which can be seen in booth 311 (*Reader Service No. 101*). Geared to reduced instruction set computing and delivering 12 million instructions per second, the \$8,995 (US) SPARCstation 1 provides, says the company, enhanced productivity for chemical database management, low-end molecular modelling, electronic publishing of technical findings and laboratory management. Special features include multitasking and windowing capabilities which allow several applications to be performed simultaneously in different windows. The SPARCstation 1 is binary compatible with all other SPARCsystem workstations and Sun-4 family machines, and will run Sun software and a range of third-party software. An optional DOS Windows emulation package is available.

World Precision Instruments, exhibiting in booth 628, will be featuring WavEdit — new **software for reviewing and editing data** recorded in Apple Macintosh files (*Reader Service No. 102*). The \$395 (US) waveform editing software can open files created by WPI data acquisition



Review and edit data with WaveEdit software from WPI.

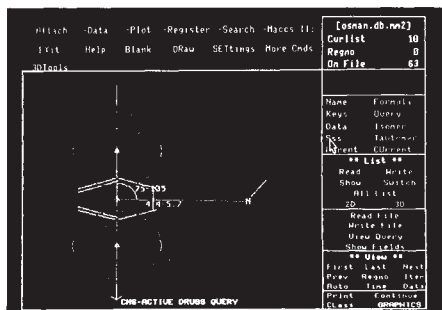
systems as well as any 12-bit data files that are saved in text or binary format. Regions of interest are selected using cursors and scrolling and indexing functions. Selections are then cut, copied and pasted into a new window. Recordings appear as chart recorder-type displays, which can be overlaid, displayed side-by-side, or compared numerically. According to WPI, the integrity of the original data is maintained as WavEdit does not delete or modify the raw data file. The software will run on any Macintosh computer that has at least one MByte of memory, however, the company says that the performance can be enhanced by using Macintosh II, IIx, or IIcx hardware. A demo version of WavEdit is available from WPI for a nominal price.

Source for Automation, Inc. will be showing off their new general purpose **robotic titration workstation** in booth 226 (*Reader Service No. 103*). Titration routines can be tailored to individual requirements by selecting from a range of sample preparation procedures, including grinding, heating, cooling, weighing, shaking, stirring, vortexing, diluting and filtering. Samples are identified by barcodes. The system then searches for the relevant titration routine, prepares the sample, and places it in the titration head, with automatic presentation of the cup to the probes. Liquids are transferred by peristaltic transfer by volume or weight, or by syringe pump or burette by volume. The workstation automatically records results and performs calculations, and allows users to create individualized report formats. The general purpose titration workstation costs about \$40,000 (US) and can function as an unattended stand-alone system or as part of a network.

Molecular modelling

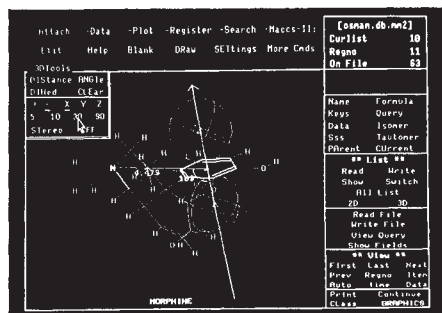
Visitors to Molecular Design Ltd's booth, number 112, will be able to see the latest addition to the graphics-based MACCS II **chemical information management system** — MACCS-3D (*Reader Service No. 104*). The software can be used for storing, searching and retrieving 3-D chemical

3. Weber, I.T. *et al. Science* **243**, 928 (1989).
4. Karplus, M. & McCammon, J.A. *Ann. Rev. Biochem.* **53**, 263 (1983).
5. McCammon, J.A. & Harvey, S.C. *Dynamics of Proteins and Nucleic Acids*, (Cambridge University Press, Cambridge, 1987).
6. Arthos, J. *et al. Cell* **57**, 469 (1989).
7. Maddon, P.J. *et al. Proc. natn Acad. Sci. U.S.A.* **84**, 9155 (1987).
8. Jameson, B.A. *et al. Science* **240**, 1335 (1988).
9. Mizukami, T., Fuerst, T.R., Berger, E.A. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **85**, 927 (1988).
10. Clayton, L.K. *et al. Nature* **335**, 363 (1988).
11. Peterson, A. & Seed, B. *Cell* **54**, 65 (1988).
12. Wernicke, D., Weiner, D. & Jameson, B.A. (in preparation).
13. Epp, O. *et al. Biochemistry* **14**, 4943 (1975).



Researchers can add geometric constraints to a query using MACCS-3 D from Molecular Design.

models and model-related data, including data specific to atoms and atom pairs. Developed by a consortium of nine chemical and pharmaceutical companies, the models in MACCS-3D are standard MACCS II structures for which additional 3-D information is stored, including Cartesian coordinates, partial atomic charges and MM2 energy. The software allows exact-match, substructure, and submodel



Conformational analyses can be performed using the system's 3-D tools.

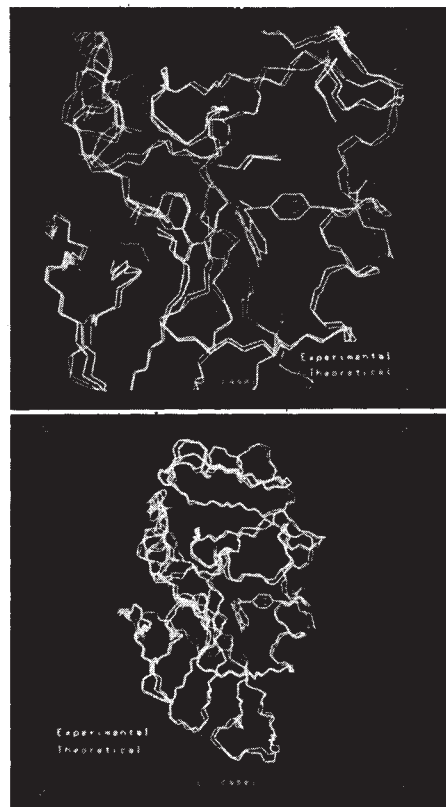
searching of 3-D models, where geometric constraints on angles, dihedral angles and exclusion spheres can either be set by the user or left at MACCS-3D's default. Molecular Design emphasizes the application of MACCS-3D in identifying candidate drugs based on geometric properties. The software will be available in November, says the company.

A second generation version of TOPFRAG — Derwent Publications' chemical fragment coding system — which allows the user to **develop search strategies for generic structures** or groups/families of chemical compounds is being developed by Hampden Data Services Ltd (*Reader Service No. 105*). Generic TOPFRAG provides automatic generation of fragmentation codes from graphic search queries of specific compounds, and produces a search strategy that aids in the search of the World Patents Index files on three major online hosts. Other features include improved drawing capabilities made possible by Hampden's Psidom industry standard Chemical Structure Drawing Interface, a wider range of PCs, printers, and mice. Generic TOPFRAG is currently undergoing β testing and is

scheduled for release at the end of the year, says the company, who can be seen in booth 425.

Chemical Design, Inc. and ORAC Ltd have joined forces to produce KERNIX — a computerized system for **chemical information management**, which is due for release later this year (*Reader Service No. 106*). Based on compound document architecture (CDA) and run under DECwindows, KERNIX is designed for the management and exchange of information between databases, computerized documents, modelling systems and other information sources. CDA uses the file format called Digital Document Interchange Format to allow files with values from databases, text, digitized images and graphics to be created, stored and exchanged over a range of devices and operating environments. Chemical Design says the use of DECwindows provides an easy-to-use windowing interface with pull-down and pop-up menus, while taking full advantage of a networking facility. Chemical Design will be in booth 322.

BIOGRAF backgrounder is the new **molecular modelling and simulation software** package from BioDesign Inc., which allows researchers to evaluate the structure, properties and behaviour of molecules at the atomic level (*Reader Service No. 107*). The software uses molecular dynamics and mechanics algorithms and conformational searches to allow re-



Two views of dihydrofolate reductase with methotrexate bound, simulated using BioDesign's BIOGRAF software.

searches to visualize and model molecules, examine their active binding sites in 3D, and perform energy calculations on the model. BIOGRAF offers support for many classes of compounds and has applications in protein research, drug design and interaction studies, and DNA modelling, says the company. Central to the accuracy of the simulation are the force fields which are based on theoretical calculations and predictions. BIOGRAF provides its own proprietary force field called Drieding but can interface with MM2, AMBER and CHARMM force fields. The software will run on selected leading computers, graphics terminals and workstations and prices range from \$30,000 to \$150,000 (US). BioDesign will be in booth 122.

Data handling

Hewlett-Packard Company, exhibiting in booth 203, will be launching HP LAB/UX — a **laboratory information management system** for the efficient organization of laboratory data (*Reader Service No. 108*). The system is based on proven LIMS software and the latest HP precision architecture computer systems. LAB/UX can be customized to suit user and laboratory needs, and can support up to 20 simultaneous users. LAB/UX can be expanded

ADVERTISEMENTS

Custom DNA

Purified and Delivered in 48 hours

100 ug, \$10.00 a base for the first 15 bases, \$7.50 for each additional base. Milligram quantities available.

Research Genetics
2130 S. Memorial Parkway Huntsville, AL 35801

1-800-533-4363
In the United Kingdom 0-800-89-1393
In Canada 1-800-833-8348
FAX 205-536-9016 (NW2971)D

Reader Service No. 15

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS — EXTREMELY RAPID

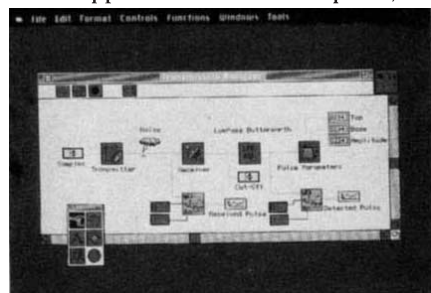
Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

BIAFORM
POB 126, D-8000 Munich 65, FRG

Reader Service No. 45

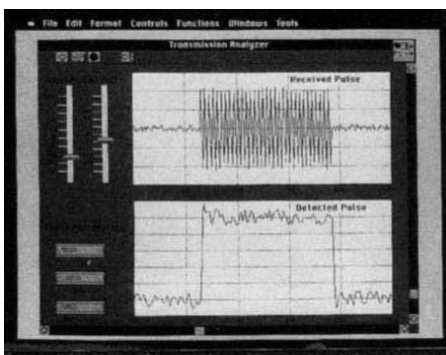
and updated without the need for reprogramming or system shut-down, says the company. Special features include sample log-in, test scheduling, result entry and validation, reporting, storage, search and retrieval of results from the database, and laboratory resource scheduling. Using the report generators, users can customize their own reports, and integrate both text and graphics into the report. Audit Trail and Labquest programs for tracking data history and for the rapid search, sort and presentation of information in the database are supplied as standard. As part of HP's Unified Laboratory Strategy, LAB/UX can communicate with PCs, mini- and main-frame computers.

National Instruments will be highlighting, in booth 519, LabVIEW 2—a **graphical programming language** that allows researchers to develop control and monitoring programs for laboratory automation, data acquisition and instrumentation (*Reader Service No. 109*). Designed to run on an Apple Macintosh II computer, the



Improved editing facilities are a feature of the Version 2 of LabVIEW from National Instruments.

software allows the build-up of control programs from functional blocks called virtual instruments (VIs). The VIs can control analysis equipment, process data or display information and, says the company, can be combined in a hierarchy to produce complex application programs. An important feature of the \$1,995 (US) LabVIEW 2 is the compiler, which con-

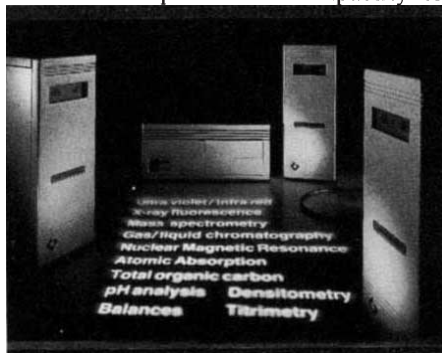


Data presentation with LabVIEW 2.

verts block diagrams into machine code, greatly increasing the execution speed, says the company. Improved editing features include wire-stretching and cut-and-paste facilities, colour data presentation, and an extended library of functions.

LabVIEW 2's open architecture allows its programs to be interfaced with those written in conventional languages.

VG Laboratory Systems, exhibiting in booth 420, has a new **data acquisition device**—the VG Instrument Server—which provides the hardware interface link between Data Manager and analytical instruments (*Reader Service No. 110*). The module provides the capability to



The VG Instrument Server.

automatically collect test results for storage in the LIMS database, speeding up data entry and eliminating possible transcription errors, says the company. Either four or eight RS232 serial lines can be attached to one instrument server, which can then be connected either directly to a VAX or via Ethernet. The Instrument Server can connect to any instrument with serial, parallel, BCD, or analog interfaces. Other features include 1/2 MByte RAM, worksheets for instruments with autosamplers, selection of data from reports, non-stop data acquisition, barcode option, and hand-held data collection.

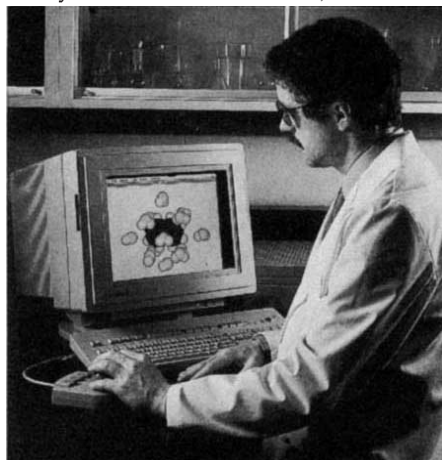
OMEGA 235 is the new **liquid chromatography software package** from Perkin-Elmer Nelson Systems for the acquisition and analysis of spectral data (*Reader Service No. 111*). Designed for use with OMEGA-2 or OMEGA-4 chromatography workstations integrated with the Perkin-Elmer LC-35 detector, the application software package provides automatic uploading from the detector after each chromatographic run. The chromatogram and spectral data can be displayed side-by-side on the same screen, stored on the system's hard disk, or printed on a GP-100 printer/plotter. The facility for multiple spectral overlays allows the user to identify peaks and confirm peak purity. When using OMEGA 235 software, all normal data handling functions of the workstation are maintained. The package comprises the OMEGA add-on LC application software and LC-235 spectral view program. Perkin-Elmer Nelson Systems will be exhibiting in booth 304.

In addition...

Designed to make conventional integrators obsolete in terms of performance and

cost, the latest Peakmaster PC-based **chromatography data station** from Harley Systems Ltd can store raw and processed data on disc, allowing reintegrations to be performed without the need for reinjection (*Reader Service No. 112*). The £2,850 (UK) system comprises a Masterlink interface and the software to connect two chromatographs to an IBM PC or compatible: a further two channels can be added for £850 (UK). Other data analysis features include the facility for the enlargement of sections of chromatograms and the subtraction of one curve from another to highlight sample differences. The Peakmaster system provides control of integration parameters, including baseline incline, convolution, skim type and skim ratio, control of external devices such as fraction collectors, pumps and autosamplers, and numerical and graphical reprocessing of chromatograms. The system can be interfaced with most chromatography instruments and can function in a stand-alone form or can be networked.

Apple Macintosh Cx or IIX users can now upgrade their machine with the CAChe Worksystem—the **stereo 3-D molecular modelling system**—for less than \$20,000 (US), says Tektronix, Inc. (*Reader Service No. 113*). The upgrade package includes an 88-K accelerator board, Trackball input device, stereo-ready 16-inch colour monitor, stereo 3-D



The CAChe Worksystem from Tektronix.

graphics card, graphics software, molecular editor software, manuals, system assembly, test and installation. The package incorporates MM2 force fields, conformational analysis, and extended Hückel for the quantum mechanical calculation of electronic properties. The CAChe Worksystem features a cut-and-paste facility for importing molecular structures into other Macintosh applications such as ChemDraw, Microsoft Word and Excel, and provides full colour graphics output support for slides, reports and documents.

Keithley Instruments Ltd has an IOtech DAC488 **digital-to-analog converter** with built-in waveform buffer that is available

in either a two- or four-channel version (*Reader Service No. 114*). The DAC488/4 four-channel version, has four analog outputs that can be programmed for range and output voltage. Each independent channel has three voltage ranges: ± 1 V, ± 5 V and ± 10 V. Outputs can be described in either volts or binary bits from the IEEE bus. Special features of the unit include an autorange feature that automatically switches to the range that provides the best resolution for the commanded voltage, and a synchronization



The DAC488: Keithley Instruments' digital-to-analog converter.

feature that allows all four channels of the digital output to be updated simultaneously. A 1024-point voltage waveform can be saved in non-volatile RAM for each output channel, which, says the company, allows the user to store a different waveform for each channel.

Map drawing

According to GeneSystems, Plasmid Artist is the first graphics program that produces publication-quality **plasmid diagrams** (*Reader Service No. 115*). Using a MacDraw-like user interface, pull-down menus and a mouse, users can draw linear or circular restriction map diagrams of recombinant DNA molecules containing, says GeneSystems, a virtually unlimited number of restriction sites and fragments. The 'Site' menu has over 60 common restriction enzyme names, but others can be added by the user. Plasmid Artist, costing about \$395 (US), automatically keeps track of the map size and positions all restriction fragments and sites in their scaled positions. GeneSystems says, high-quality images can be printed on a PostScript printer. Alternatively, the image can be exported as an encapsulated PostScript file to page layout programs such as Page Maker and Cricket Draw for incorporation into reports and manuscripts.

CLONE 2 is the upgraded version of the CLONE Manager Program from Scientific & Educational Software for **constructing and drawing circular and linear restriction maps** of recombinant DNA molecules (*Reader Service No. 116*). Designed to run on an IBM PC/XT, AT, PS/2, CLONE 2 uses pull-down menus to create maps of presentation quality, says the company. Enzymes can be used to 'cut' molecules

into fragments, which can then be ligated onto other DNA fragments. After ligation, the base positions of enzyme sites, genes and markers are automatically recalculated and the maps can be printed immediately. The program automatically determines the cohesive ends produced by each enzyme used, and uses this information to identify compatible ends. CLONE 2 contains a comprehensive list of restriction enzymes, however, the user may add to these if necessary. The program is supplied with a 50-page user manual: demo versions are available on request.

Handy helpers

Using only a few keystrokes, users can **convert units of measure** from one unit system to another with the S.I. Plus PC software program from Geocomp Corp. (*Reader Service No. 117*). Designed to run on IBM PCs or compatibles, S.I. Plus contains all the unit conversion factors commonly used by engineers and scientists, and over 70,000 different conversions can be performed in 80 classes of units, says Geocomp. The \$79 (US) S.I. Plus can also be used as a memory-resident program, whereby the S.I. Plus menu can be accessed while running another program.

Version 4.2 of Get-A-Ref — DatAid, Inc.'s **reference handling software program** — is now available (*Reader Service No. 118*). Get-A-Ref is designed to run on IBM PC/XT/AT/386/PS computers or compatibles and features an Editor with context-sensitive help, an improved user interface, and a cut-and-paste facility for integrating citations into manuscripts. The software comes complete with its own database conversion utility, which allows users of databases such as Medline, BIOSIS, Chemical Abstracts, Excerpta Medica and IRIS to convert screen dumps of information into Get-A-Ref files. According to DatAid, the program can store up to 32,000 references in ASCII format, with up to 16,000 characters of text in each file. Searches can be conducted at a rate of 100,000 characters per second and without the need for user-defined keywords, says the company.

Mathematical problem solving is made easy with the Version 4.3 of Maple — an interactive **computer algebra system** for both symbolic and numerical computations — from Waterloo Maple Software (*Reader Service No. 119*). Designed for IBM PS/2 Models 70 and 80, and other 80386-based machines running PC-DOS or MS-DOS, the \$1,100 (Cdn) Maple software has scientific, engineering, business and educational applications. The software uses standard mathematical notation for problem solving in an interactive environment, and can be used to solve equations and perform integrals at all levels. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain additional information about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

ADVERTISEMENTS

Teflon^{*} Coated and Bare Wire from MEDWIRE



Medwire provides the highest quality platinum, platinum alloy, silver and stainless steel wire in both single and multistrand configurations.

For prompt delivery or additional information, please contact:

MEDWIRE, an affiliate of Sigmund Cohn Corp.
121 So. Columbus Ave. Mount Vernon, NY 10553
Tel: (914) 664-5300 Fax: (914) 664-5377

*Reg. TM of E.I. DuPont & Co.

SINGER INSTRUMENTS COMPLETE SYSTEMS FOR



XENOPUS OOCYTE INJECTION

INT. SALES: +44. 984. 40226.

Singer Instrument Co. Ltd.

Roadwater, Watchet, Somerset,
England, TA23 0QL. Tel: Washford (0984) 40226.
Telex: 337492 Comcab G

ACTIVITY MONITOR

Automatic measurement of movement from the video (TV) image. For water maze (pool), open field, radial maze etc. IBM-PC compatible. Software for latency, path length, percent time moving, position preference etc. Options for rearing and social interaction.

In the U.S. call 1-800-225 9261

HVS IMAGE

18 Ormond Crescent
Hampton TW12 2TH
England
Telex: 8813487 LCLAD
Tel: +44 1 783 1152



See us at Neuroscience,
Phoenix, Oct. 30 - Nov. 2

Reader Service No.25

Reader Service No.50

Reader Service No.18

Information skills still in demand

Richard Pearson

Employers may have to pay a high price for not taking the responsibility for training their own staff in information technology skills.

THE production and use of information technology (IT) continues to be a key factor in the growth of many economies. In the United Kingdom the market has been growing at around 15 per cent a year, and is now worth more than £20,000 million. But since the early 1980s, there has been concern that shortages of suitably qualified staff have been constraining the development of both the IT industry and the use of IT throughout the economy.

The fastest rates of IT employment growth in the late 1980s have been in the IT consultancy, software and services companies. In electronics, employment has been recovering after the slump of 1986-87. Among the users, growth has been variable, but has been strongest in the service sector with industry still lagging behind'. In 1988, there were more than 240,000 professional IT staff employed in the United Kingdom, an increase of about 5 per cent since 1987, and 20 per cent since 1985. Total employment is now divided almost equally between the electronics and IT services sectors and users of IT in industry and the service sector; this represents a shift towards the user sector.

The recruitment of experienced staff continues to be the principal way in which most employers fill vacancies for professional IT staff. This, however, simply involves a reallocation of existing staff between organizations as employers chase staff with two or more years' experience of a particular technology or application. New graduates remain the main source of new entrants to the stock of professional IT staff, with demand continuing to rise throughout the 1980s.

In 1988, half the graduates recruited for IT work had electronics or computer studies degrees, while most of the remainder were from numerate scientific or technical disciplines. Electronics graduates went chiefly to the electronics sector. Computer science graduates were recruited to all sectors although recruiters still have mixed views about the quality and relevance of such graduates. Graduates from non-IT disciplines were principally recruited for computing work, few being taken into electronics areas. Very few graduates from non-numerate disciplines take up computing work. There remains only a small demand for postgraduate qualifications in IT, principally for research work in the electronics sector and higher education.

The output of IT graduates increased

rapidly in the 1980s, and most electronics and computer science graduates moved straight into relevant employment. The downturn in female students both applying for and entering university IT degree courses has been of particular concern, although some improvement is visible in the 1988 figures. The fall-off in interest from women has been particularly marked in computer studies; it is also reflected in the difficulties of many employers in their attempts to increase the recruitment of women to IT posts.

Recruitment difficulties continue across all professional IT occupations and sectors; the improvement shown in 1987 has not continued into 1988. Problems have been most acute for employers seeking experienced computing staff and software engineers and project managers/leaders, and where very specific technical expertise is required or combinations of technical/commercial/personal skills. Problems in recruiting new graduates seem to have been least acute in South-East England.

An increasing but still small number of employers are responding to recruitment difficulties by improving career development, retraining and deployment of staff; seeking to control wastage; increasing internal pay flexibility; reorganizing and redesigning IT work; and in some cases relocating the work to easier provincial labour markets. But progress remains slow outside the public and financial sectors where there is a history of internal staff development and retraining and salary increases remain the primary response for most employers.

Looking to the future, continued growth is expected in the IT market and in employers' needs for professional IT staff to service both business growth and the growing use of IT in new areas. Considerable uncertainty exists, however, about the rate of market growth, and the effects on employment patterns of mergers and acquisitions, European markets and competition and the contracting out of IT work.

Few employers have a developed form of manpower planning. Most plan employment and recruitment levels over no more than 12-18 months and are therefore unable to forecast future IT employment levels with any precision.

What is clear, however, is that the labour market for professional IT staff will become increasingly compartmentalized,

with the balance of supply and demand varying between sectors, locations, occupations and types of experience. There is expected to be a continuing growth in the demand for software and computing skills and applications experience, but little growth in the case of hardware skills. With the growing 'user-friendliness' and spread of IT, there will be a need for more 'hybrid' staff with both IT and application skills. Occupational boundaries will become increasingly blurred. The main growth occupations are likely to be based on communication and networks and IT consultants/business analysts because of the greater integration of IT into the business. Demand for analysts/programmers and computing is also expected to continue to grow. Overall demand is likely to grow at under 5 per cent per annum, assuming economic growth continues, representing a slowdown on recent years.

Graduates will continue to be the main source of new entrants to IT jobs and graduate intakes are expected to increase. While the supply of IT graduates is expected to continue to rise into the early 1990s, the increased demand from recruiters in general will put pressure on recruiters seeking to fill IT jobs from a wider range of subjects than just IT. In the mid-1990s, the output of graduates is expected to stop growing, because of the demographic downturn'. As demand is expected to continue to rise, there will be increasing competition from IT and other employers for graduates from IT and other numerate disciplines.

The outlook is for a continuation of the tight labour market for IT skilled staff. Employers will continue to face difficulties recruiting experienced professional IT staff unless they give more attention to the use of in-company resources for retraining/conversion, to non-traditional recruitment sources and to greater flexibility in working arrangements. A start has been made by a few employers, but there is need for more innovative practices both within the IT industry and beyond in the expanding user sector. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton, Sussex, BN1 9RF, UK.

1. Connor, H., Pearson, R. & Buchan, J. *The Changing IT Skills Scene: IT Manpower Monitor 1989* (IMS, Brighton, 1989).
2. Pearson, R. & Pike, G. *The Graduate Labour Market in the 1990s* (IMS, Brighton, 1989).